Guía de publicación de Amazon EMR

Amazon EMR

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

Amazon EMR: Guía de publicación de Amazon EMR

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

Las marcas comerciales y la imagen comercial de Amazon no se pueden utilizar en relación con ningún producto o servicio que no sea de Amazon, de ninguna manera que pueda causar confusión entre los clientes y que menosprecie o desacredite a Amazon. Todas las demás marcas registradas que no son propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, conectados o patrocinados por Amazon.

Table of Contents

Acerca de las versiones de Amazon EMR 1
Soporte estándar 2
Descripción de las versiones de Amazon EMR 2
Control de versiones de lanzamiento 3
Política de soporte 4
Consideraciones 8
Versiones de Amazon EMR 7.x 9
Versiones de las aplicaciones en las versiones 7.x de Amazon EMR 10
emr-7.9.0 10
emr-7.8.0 50
emr-7.7.0 97
emr-7.6.0 151
emr-7.5.0 230
emr-7.4.0 307
emr-7.3.0 386
emr-7.2.0 465
emr-7.1.0 584
emr-7.0.0 644
Versiones de lanzamiento de Amazon EMR 6.x 714
Versiones de la aplicación en las versiones 6.x de Amazon EMR 715
emr-6.15.0 716
emr-6.14.0 784
emr-6.13.0 848
emr-6.12.0 923
emr-6.11.1 1003
emr-6.11.0 1079
emr-6.10.1 1137
emr-6.10.0 1213
emr-6.9.1 1275
emr-6.9.0 1341
emr-6.8.1 1407
emr-6.8.0 1472
emr-6.7.0 1536
emr-6.6.0 1612

emr-6.5.0 1690
emr-6.4.0 1719
emr-6.3.1 1754
emr-6.3.0 1782
emr-6.2.1 1816
emr-6.2.0 1844
emr-6.1.1 1879
emr-6.1.0 1900
emr-6.0.1 1927
emr-6.0.0 1945
Versiones de lanzamiento de Amazon EMR 5.x 1967
Versiones de la aplicación en las versiones 5.x de Amazon EMR 1971
emr-5.36.2 1971
emr-5.36.1 2018
emr-5.36.0 2082
emr-5.35.0 2132
emr-5.34.0 2160
emr-5.33.1 2187
emr-5.33.0 2217
emr-5.32.1 2240
emr-5.32.0 2265
emr-5.31.1 2294
emr-5.31.0 2314
emr-5.30.2 2338
emr-5.30.1 2358
emr-5.30.0 2381
emr-5.29.0 2404
emr-5.28.1 2423
emr-5.28.0 2442
emr-5.27.1 2463
emr-5.27.0 2480
emr-5.26.0 2500
emr-5.25.0 2520
emr-5.24.1 2540
emr-5.24.0 2558
emr-5.23.1 2577

emr-5.23.0 2594
emr-5.22.0 2613
emr-5.21.2 2633
emr-5.21.1 2650
emr-5.21.0 2668
emr-5.20.1 2688
emr-5.20.0 2705
emr-5.19.1 2725
emr-5.19.0 2742
emr-5.18.1 2760
emr-5.18.0 2777
emr-5.17.2 2795
emr-5.17.1 2812
emr-5.17.0 2829
emr-5.16.1 2846
emr-5.16.0 2863
emr-5.15.1 2880
emr-5.15.0 2896
emr-5.14.2 2914
emr-5.14.1 2930
emr-5.14.0 2946
emr-5.13.1 2965
emr-5.13.0 2980
emr-5.12.3 2997
emr-5.12.2 3012
emr-5.12.1 3028
emr-5.12.0 3044
emr-5.11.4 3061
emr-5.11.3 3077
emr-5.11.2 3093
emr-5.11.1 3108
emr-5.11.0 3124
emr-5.10.1 3140
emr-5.10.0 3156
emr-5.9.1 3173
emr-5.9.0 3189

emr-5.8.3 3206
emr-5.8.2 3221
emr-5.8.1 3236
emr-5.8.0 3251
emr-5.7.1 3268
emr-5.7.0 3283
emr-5.6.1 3298
emr-5.6.0 3314
emr-5.5.4 3329
emr-5.5.3 3344
emr-5.5.2 3359
emr-5.5.1 3374
emr-5.5.0 3389
emr-5.4.1 3406
emr-5.4.0 3421
emr-5.3.2 3436
emr-5.3.1 3451
emr-5.3.0 3466
emr-5.2.3 3481
emr-5.2.2 3496
emr-5.2.1 3511
emr-5.2.0 3526
emr-5.1.1 3542
emr-5.1.0 3556
emr-5.0.3 3571
emr-5.0.2 3586
emr-5.0.1 3600
emr-5.0.0 3614
Versiones de lanzamiento de Amazon EMR 4.x 3629
Versiones de la aplicación en las versiones 4.x de Amazon EMR 3631
Diferencias entre las versiones de lanzamiento 3631
emr-4.9.6 3679
emr-4.9.5 3694
emr-4.9.4 3708
emr-4.9.3 3723
emr-4.9.2 3737

emr-4.9.1 3751
emr-4.8.5 3766
emr-4.8.4 3780
emr-4.8.3 3795
emr-4.8.2 3810
emr-4.8.1 3824
emr-4.8.0 3839
emr-4.7.4 3854
emr-4.7.3 3867
emr-4.7.2 3880
emr-4.7.1 3894
emr-4.7.0 3908
emr-4.6.1 3923
emr-4.6.0 3935
emr-4.5.0 3949
emr-4.4.0 3960
emr-4.3.0 3973
emr-4.2.0 3984
emr-4.1.0 3995
emr-4.0.0 4004
Versiones 2.x y 3.x de la AMI 4011
Creación de un clúster 4013
Instalación de aplicaciones 4015
Personalización de configuraciones 4015
Hive 4023
HBase 4034
Pig 4047
Spark 4053
S3 DistCp 4057
Novedades 4060
Amazon EMR 7.9.0 4060
Amazon EMR 6.15.0 4071
Amazon EMR 5.36.2 4108
Compatibilidad con el SigV4 4130
Método para mitigar CVE-2021-44228 4130

Solución de acción de arranque de Amazon EMR para Log4j CVE-2021-44228 y CVE-2021-45046 ... 4131
Preguntas frecuentes ... 4139
Archivado ... 4141
Versión 6.14.0 ... 4142
Versión 6.13.0 ... 4177
Versión 6.12.0 ... 4221
Versión 6.11.1 ... 4270
Versión 6.11.0 ... 4316
Versión 6.10.0 ... 4346
Versión 6.9.0 ... 4379
Versión 6.8.0 ... 4415
Versión 6.7.0 ... 4452
Versión 6.6.0 ... 4500
Versión 5.35.0 ... 4551
Versión 5.34.0 ... 4555
Versión 6.5.0 ... 4558
Versión 6.4.0 ... 4560
Versión 5.32.0 ... 4568
Versión 6.2.0 ... 4574
Versión 5.31.0 ... 4582
Versión 6.1.0 ... 4588
Versión 6.0.0 ... 4594
Versión 5.30.1 ... 4600
Versión 5.30.0 ... 4605
Versión 5.29.0 ... 4610
Versión 5.28.1 ... 4612
Versión 5.28.0 ... 4613
Versión 5.27.0 ... 4615
Versión 5.26.0 ... 4618
Versión 5.25.0 ... 4621
Versión 5.24.1 ... 4624
Versión 5.24.0 ... 4625
Versión 5.23.0 ... 4627
Versión 5.22.0 ... 4630
Versión 5.21.1 ... 4633

Versión 5.21.0 4634
Versión 5.20.0 4637
Versión 5.19.0 4640
Versión 5.18.0 4642
Versión 5.17.1 4643
Versión 5.17.0 4644
Versión 5.16.0 4645
Versión 5.15.0 4646
Versión 5.14.1 4647
Versión 5.14.0 4648
Versión 5.13.0 4650
Versión 5.12.2 4651
Versión 5.12.1 4651
Versión 5.12.0 4651
Versión 5.11.3 4653
Versión 5.11.2 4653
Versión 5.11.1 4654
Versión 5.11.0 4654
Versión 5.10.0 4655
Versión 5.9.0 4657
Versión 5.8.2 4659
Versión 5.8.1 4660
Versión 5.8.0 4660
Versión 5.7.0 4662
Versión 5.6.0 4663
Versión 5.5.3 4664
Versión 5.5.2 4664
Versión 5.5.1 4664
Versión 5.5.0 4665
Versión 5.4.0 4666
Versión 5.3.1 4667
Versión 5.3.0 4667
Versión 5.2.2 4668
Versión 5.2.1 4669
Versión 5.2.0 4670
Versión 5.1.0 4670

Versión 5.0.3 4671
Versión 5.0.0 4672
Versión 4.9.5 4673
Versión 4.9.4 4674
Versión 4.9.3 4674
Versión 4.9.2 4674
Versión 4.9.1 4674
Versión 4.8.4 4675
Versión 4.8.3 4675
Versión 4.8.2 4676
Versión 4.8.0 4677
Versión 4.7.2 4677
Versión 4.7.1 4678
Versión 4.7.0 4679
Versión 4.6.0 4680
Versión 4.5.0 4682
Versión 4.4.0 4683
Versión 4.3.0 4685
Versión 4.2.0 4687
Configuración de aplicaciones 4688
Configuración de las aplicaciones al crear un clúster 4691
Proporcionar una configuración en la consola al crear un clúster 4691
Proporcione una configuración mediante AWS CLI al crear un clúster 4691
Proporcione una configuración mediante el SDK de Java al crear un clúster 4692
Reconfiguración de un grupo de instancias en un clúster en ejecución 4693
Consideraciones a la hora de volver a configurar un grupo de instancias 4693
Reconfiguración de un grupo de instancias en la consola 4697
Reconfiguración de un grupo de instancias con la CLI 4698
Reconfiguración de un grupo de instancias con el SDK para Java 4703
Solución de problemas 4704
Almacenamiento de datos de configuración confidenciales en AWS Secrets Manager 4707
Creación de un secreto 4707
Concesión de acceso a Amazon EMR para recuperar el secreto 4707
Uso del secreto en una clasificación de configuración 4708
Actualización del valor del secreto 4709
Configuración de aplicaciones para utilizar una máquina virtual de Java específica 4709

Consideraciones 4709
Anulación de la JVM 4712
Puertos de servicio 4715
Usuarios de la aplicación 4717
Comprobación de dependencias mediante el repositorio de artefactos 4719
Sistema de archivos de EMR (EMRFS) 4722
Consistent view 4724
Habilitación de la vista coherente 4728
Seguimiento de objetos de EMRFS mediante la vista consistente de Amazon S3 4730
Lógica de reintentos 4731
Metadatos de la vista consistente de EMRFS 4733
Configurar notificaciones de consistencia para Amazon CloudWatch SQS y Amazon SQS 4736
Configuración de la vista coherente 4738
La Referencia de los comandos de la CLI de EMRFS 4743
Autorización de acceso a los datos de EMRFS en Amazon S3 4755
Creación de un proveedor de credenciales personalizado para los datos de EMRFS en Amazon S3 4755
Administrar el AWS Security Token Service punto de conexión de predeterminado 4757
Especificación del cifrado de Amazon S3 mediante propiedades de EMRFS 4758
Utilización de AWS KMS keys para el cifrado de EMRFS 4759
Cifrado del servidor de Amazon S3 4760
Cifrado del cliente de Amazon S3 4763
Cifrado del lado del cliente de Amazon S3 con S3A 4772
Configuración CSE-KMS 4772
Configuración: CSE-CUSTOM 4773
Propiedades del cifrado del lado del cliente de Amazon S3 con S3A 4774
CloudWatch Agente de Amazon 4777
Creación de un clúster 4779
Permisos necesarios 4779
Punto de conexión requerido 4780
Cree el clúster 4780
Métricas predeterminadas 4781
Configuración 4784
Amazon EMR 7.1.0 4784
Amazon EMR 7.0.0 4820
Consideraciones 4826

Historial 4826
Delta Lake 4832
Introducción 4832
Uso de clústeres de Delta Lake 4833
Delta Lake con Flink 4833
Delta Lake con Trino 4838
Delta Lake con Spark 4839
Delta Lake con Spark y Glue 4846
Consideraciones 4846
Historial 4847
Flink 4850
Creación de un clúster con Flink 4853
Configuración de Flink 4854
Hive y Glue 4854
Archivo de configuración 4856
Múltiples nodos principales 4857
Tamaño del proceso de memoria 4858
Tamaño del archivo de salida del registro 4859
Java 11 4860
Trabajos de Flink 4865
Inicio de una aplicación de YARN de Flink como un paso en un clúster de ejecución prolongada 4866
Envío del trabajo a una aplicación de Flink existente en un clúster de ejecución prolongada 4867
Enviar un trabajo de Flink transitorio 4868
Intérprete de comandos de Scala para Flink 4871
Interfaz de usuario de Flink 4871
Escalador automático de Flink 4873
Descripción general 4873
Consideraciones 4873
Habilitación del escalador automático 4874
Configuraciones 4875
Optimización de los tiempos de reinicio 4882
Recuperación local de tareas 4882
Puntos de control incrementales 4883
Recuperación detallada 4884

Reinicio combinado 4885
Flink con Zeppelin 4886
Introducción 4886
Requisitos previos 4886
Configuración de Zeppelin-Flink en un clúster de EMR 4887
Ejecución de trabajos de Flink con Zeppelin-Flink en un clúster de EMR 4888
Historial de lanzamientos de Flink 4893
Notas de la versión de Flink por versión 4934
Ganglia 4937
Creación de un clúster con Ganglia 4939
Ver métricas de Ganglia 4940
Métricas de Hadoop y Spark en Ganglia 4941
Historial de versiones de Ganglia 4942
Hadoop 4988
Configurar Hadoop 4990
Configuración de tareas 4991
Ajustes de configuración del daemon de Hadoop 5425
Configuración de HDFS 5751
Cifrado transparente en el HDFS en Amazon EMR 5752
Configuración del cifrado transparente de HDFS 5753
Consideraciones para cifrado transparente de HDFS 5756
Hadoop Key Management Server 5756
Cifrado transparente de HDFS en clústeres de EMR con varios nodos maestros 5760
Crear o ejecutar una aplicación de Hadoop 5762
Creación de binarios con Amazon EMR 5763
Procesamiento de datos con streaming 5765
Procesamiento de datos con un JAR personalizado 5771
Leer objetos restaurados 5775
Ejemplos 5775
Consideraciones 5776
Activar el reconocimiento de acceso no uniforme a la memoria para los contenedores YARN 5776
Embalaje de contenedores YARN 5779
Habilite la función 5779
Consideraciones 5779
Historial de versiones de Hadoop 5780
Notas de la versión de Hadoop por versión 5845

HBase 5859
Crear un clúster con HBase 5863
Crear un clúster HBase mediante la consola 5863
Creación de un clúster mediante el uso de HBase AWS CLI 5863
HBase en Amazon S3 (modo de almacenamiento de Amazon S3) 5865
Habilitación HBase en Amazon S3 5866
Uso de un clúster de réplicas de lectura 5867
Seguimiento persistente HFile 5869
Seguimiento de archivos de la tienda 5871
Consideraciones operativas 5872
Write Ahead Logs (WAL) para Amazon EMR 5877
Espacios de trabajo de WAL 5878
Permisos necesarios 5879
Habilitación de WAL 5880
Restauración desde WAL 5883
Configuraciones de seguridad 5884
Usando AWS PrivateLink 5885
Precios y métricas de WAL 5886
Etiquetado de espacios de trabajo de WAL 5888
Replicación entre clústeres EMR WAL 5889
Consideraciones y disponibilidad 5891
Referencia de la CLI de EMRWAL 5893
Uso del HBase shell 5896
Creación de una tabla 5897
Colocar un valor 5897
Obtener un valor 5897
Eliminación de una tabla 5897
Acceda a HBase las tablas con Hive 5897
Uso de instantáneas HBase 5899
Crear una instantánea con una tabla 5899
Eliminar una instantánea 5900
Ver información de instantánea 5900
Exportar una instantánea a Amazon S3 5900
Importar instantáneas desde Amazon S3 5901
Restaurar una tabla a partir de instantáneas del shell HBase 5902
Configuración HBase 5903

Cambios en la asignación de memoria en YARN 5904
HBase números de puerto 5904
HBase configuración del sitio para optimizar 5905
Consideraciones sobre el rendimiento 5907
Vea la interfaz de HBase usuario 5909
Ver archivos de HBase registro 5911
Supervise HBase con Ganglia 5912
Supervisión de EMR con Amazon HBase CloudWatch 5914
Configura las métricas 5916
Uso del destino de las métricas 5918
¿Estás migrando desde versiones anteriores HBase 5919
Migración a Amazon EMR versión 7.4.0 o posterior 5920
HBase historial de versiones 5921
HCatalog 6000
Crear un clúster con HCatalog 6001
Usando HCatalog 6002
Deshabilitación de la escritura directa al utilizar HCatalog HStorer 6003
Creación de una tabla mediante la HCat CLI y uso de dichos datos en Pig 6003
Acceso a la tabla mediante Spark SQL 6004
Ejemplo: crear una HCatalog tabla y escribir en ella utilizando Pig 6006
HCatalog historial de lanzamientos 6007
Hive 6079
Diferencias y consideraciones sobre Hive en Amazon EMR 6082
Diferencias entre Apache Hive en Amazon EMR y Apache Hive 6082
Diferencias de Hive entre la versión de lanzamiento 4.x y 5.x de Amazon EMR 6083
Características adicionales de Hive en Amazon EMR 6084
Configuración de un metaalmacén externo para Hive 6089
Se ha agregado compatibilidad para utilizar el Catálogo de datos de AWS Glue como metaalmacén predeterminado de Hive 6090
Uso de una base de datos MySQL externa o Amazon Aurora 6097
Uso del controlador JDBC de Hive 6100
Mejora del rendimiento de Hive 6102
Habilitación del confirmador optimizado Hive EMRFS S3 6103
Uso de S3 Select 6104
Optimización de MSCK 6107
Uso de Hive LLAP 6108

Habilitación de LLAP 6108
Inicio de LLAP en su clúster 6109
Comprobación del estado de LLAP 6110
Inicio o detención de LLAP 6110
Redimensionado del número de daemons LLAP 6111
Cifrado en Hive 6111
Cifrado modular de Parquet en Hive 6111
Cifrado en tránsito en HS2 6115
Historial de versiones de Hive 6116
Notas de la versión de Hive por versión 6190
Hudi 6286
Cómo funciona Hudi 6288
Descripción de los tipos de almacenamiento de conjuntos de datos: escribir frente a combinar al leer 6288
Registro de un conjunto de datos de Hudi con su metaalmacén 6290
Consideraciones y limitaciones 6290
Crear un clúster con Hudi instalado 6292
Uso de un conjunto de datos de Hudi 6293
Iniciar una sesión de Spark para Hudi 6297
Escribir en un conjunto de datos de Hudi 6297
Datos de upsert 6303
Eliminación de un registro 6304
Leer un conjunto de datos de Hudi 6305
Usar la CLI de Hudi 6307
..... 6308
Historial de versiones de Hudi 6309
Hue 6313
Versiones de Hue 6313
Características de Hue admitidas y no admitidas en Amazon EMR 6315
Consideraciones 6316
Rendimiento con tablas de metadatos de Hue de gran tamaño 6317
Incompatibilidad entre versiones de Hue 6318
Conexión a la interfaz de usuario web de Hue 6318
Uso de Hue con una base de datos remota en Amazon RDS 6319
Solución de problemas 6321
Configuraciones avanzadas para Hue 6322

Configurar Hue para usuarios de LDAP ... 6322
Historial de versiones de Hue ... 6325
Iceberg ... 6391
Cómo funciona Iceberg ... 6392
Utilice un clúster Iceberg ... 6394
Uso de un clúster de Iceberg con Spark ... 6394
Usar un clúster de Iceberg con Trino ... 6403
Usar un clúster de Iceberg con Flink ... 6405
Usar un clúster de Iceberg con Hive ... 6410
Historial de versiones de Iceberg ... 6415
Notas de la versión de Iceberg por versión ... 6416
Cuaderno de Jupyter ... 6419
EMR Studio ... 6419
Cuaderno de EMR ... 6419
JupyterHub ... 6420
Cree un clúster con JupyterHub ... 6424
Consideraciones sobre el uso JupyterHub en Amazon EMR ... 6426
Configurando JupyterHub ... 6427
Configuración de la persistencia de los cuadernos en Amazon S3 ... 6428
Conexión al nodo maestro y a los servidores de blocs de notas ... 6429
JupyterHub configuración y administración ... 6430
Agregar usuarios y administradores de Cuadernos de Jupyter ... 6432
Instalación de kernels y bibliotecas adicionales ... 6443
JupyterHub historial de lanzamientos ... 6448
Livy ... 6488
Habilitación de HTTPS ... 6490
Historial de versiones de Livy ... 6491
MXNet ... 6539
MXNet historial de versiones ... 6541
Oozie ... 6570
Uso de Oozie con una base de datos remota en Amazon RDS ... 6573
Configurar la versión Java para Oozie ... 6575
Historial de versiones de Oozie ... 6576
Notas de la versión de Oozie por versión ... 6637
Phoenix ... 6640
Creación de un clúster con Phoenix ... 6643

Personalización de configuraciones de Phoenix ... 6644
Clientes de Phoenix ... 6645
Historial de versiones de Phoenix ... 6648
Pig ... 6730
Enviar trabajo de Pig ... 6732
Enviar trabajo de Pig mediante la consola de Amazon EMR ... 6733
Enviar trabajo de Pig mediante la AWS CLI ... 6734
Llamada a funciones definidas por el usuario desde Pig ... 6735
Llamar a archivos de JAR desde Pig ... 6736
Llamada a scripts de Python/Jython desde Pig ... 6736
Historial de versiones de Pig ... 6737
Presto ... 6806
Uso de Presto con el Catálogo de datos de AWS Glue ... 6808
Especificación del Catálogo de datos de AWS Glue como metaalmacén ... 6809
Permisos de IAM ... 6094
Consideraciones a la hora de utilizar el Catálogo de datos de AWS Glue ... 6814
Uso de S3 Select Pushdown ... 6815
¿Es S3 Select Pushdown adecuado para mi aplicación? ... 6815
Consideraciones y limitaciones ... 6816
Habilitar S3 Select Pushdown con PrestoDB o Trino ... 6816
Agregar conectores de base de datos ... 6817
Uso de SSL/TLS y LDAPS ... 6818
Uso de la autenticación LDAP ... 6819
Activación del modo estricto de Presto ... 6827
Consideraciones ... 6829
Gestión de la pérdida de instancias de spot en Presto ... 6830
Uso del escalado automático de Presto con la retirada estable ... 6832
Consideraciones relacionadas con el uso de Presto en Amazon EMR ... 6833
Ejecutable de línea de comandos de Presto ... 6833
Propiedades de la implementación de Presto no configurables ... 6833
Instalación de PrestoDB ... 6834
Configuración de EMRFS y PrestoS3PrestoS3 FileSystem ... 6834
Configuración predeterminada para la suplantación del usuario final ... 6835
Puerto predeterminado para la interfaz web de Presto ... 6836
Problema con la ejecución del bucket de Hive en algunas versiones ... 6836
Historial de versiones de Presto ... 6836

Spark 6894
Creación de un clúster de Spark 6898
Ejecución de aplicaciones de Spark con Docker en Amazon EMR 6.x 6901
Consideraciones 6901
Creación de una imagen de Docker 6902
Uso de imágenes de Docker de Amazon ECR 6903
Usa el catálogo de AWS Glue Data Catalog con Spark en Amazon EMR 6909
Especificar AWS Glue Data Catalog como metastore de Apache Hive 6910
Especificar AWS Glue Data Catalog como catálogo de Apache Iceberg 6912
Permisos de IAM 6094
Consideraciones 6095
Trabajar con una jerarquía de varios catálogos en AWS Glue Data Catalog 6915
Cómo se organizan los recursos del catálogo 6915
Configuración de un catálogo de datos para su uso con Amazon EMR 6916
Permisos para acceder a los recursos del catálogo de datos de AWS Glue 6916
Configurar Spark para acceder a una jerarquía de varios catálogos en AWS Glue Data Catalog 6917
Consideraciones y limitaciones de una configuración de varios catálogos 6920
Configurar Spark 6920
Valores predeterminados de Spark 6921
Configuración de la recopilación de elementos no utilizados 6923
`maximizeResourceAllocation` 6924
Comportamiento de desactivación 6926
Variable de entorno Spark ThriftServer 6930
Cambio de la configuración predeterminada de Spark 6931
Migración de Apache Log4j 1.x a Log4j 2.x 6933
Optimización del desempeño de Spark 6934
Ejecución de consultas adaptativas 6934
Eliminación dinámica de particiones 6936
Aplanamiento de subconsultas escalares 6939
DISTINCT antes de INTERSECT 6939
Unión de filtros de Bloom 6940
Reordenación de unión optimizada 6941
Almacenamiento en caché de fragmentos de resultados 6942
Habilitación del Almacenamiento en caché de fragmentos de resultados de Spark 6943
Consideraciones 6944

Uso de RAPIDS Accelerator ... 6946
Elección de tipos de instancias ... 6946
Definición de las configuraciones de aplicaciones ... 6946
Adición de una acción de arranque ... 6954
Lanzamiento del clúster ... 6955
Acceso al intérprete de comandos de Spark ... 6955
Utilice Amazon SageMaker Spark para el aprendizaje automático ... 6957
Escritura de una aplicación de Spark ... 6957
Scala ... 6958
Java ... 6958
Python ... 6960
Mejora del rendimiento de Spark con S3 ... 6961
Uso de S3 Select ... 6962
Confirmador optimizado para S3 de EMRFS ... 6966
Uso del protocolo de confirmación optimizado para S3 de EMRFS ... 6973
Reintento de solicitudes de S3 ... 6980
Adición de un paso de Spark ... 6983
Anulación de las opciones de configuración predeterminadas de Spark ... 6987
Visualización del historial de aplicaciones de Spark ... 6987
Accede a la web de Spark UIs ... 6987
Uso de Spark con Amazon Kinesis Data Streams ... 6988
Uso de Spark en Amazon Redshift ... 6988
Lanzar una aplicación de Spark ... 6989
Autenticarse en Amazon Redshift ... 6990
Lectura y escritura en Amazon Redshift ... 6992
Consideraciones ... 6994
Historial de lanzamientos de Spark ... 6995
Sqoop ... 7062
Versiones de Sqoop ... 7062
Consideraciones relacionadas con el uso de Sqoop en Amazon EMR ... 7064
Uso de Sqoop con integración HCatalog ... 7064
Compatibilidad con JDBC y bases de datos de Sqoop ... 7065
Cómo proteger la contraseña ... 7066
Historial de versiones de Sqoop ... 7067
TensorFlow ... 7122
TensorFlow compilaciones por tipo de EC2 instancia de Amazon ... 7124

Seguridad ... 7124
Usando TensorBoard ... 7125
TensorFlow historial de lanzamientos ... 7126
Tez ... 7150
Creación de un clúster con Tez ... 7152
Configuración de Tez ... 7153
Configuración de ejemplo ... 7153
Apertura de división asíncrona de Tez ... 7154
IU web de Tez ... 7157
Timeline Server ... 7158
Historial de versiones de Tez ... 7158
Notas de la versión de Tez por versión ... 7204
Trino ... 7215
Historia y diseño de Trino ... 7216
Conceptos arquitectónicos ... 7216
Introducción a Trino ... 7218
Complete los pasos previos para usar Amazon EMR con Trino ... 7218
Lanzar un clúster de Amazon EMR con Trino ... 7218
Conéctese al nodo principal y ejecute consultas ... 7221
Configuración de Trino en Amazon EMR ... 7222
Configuración de conectores para Trino ... 7222
Monitorización ... 7223
Prácticas recomendadas para Trino en Amazon EMR ... 7223
Áreas clave de enfoque para la mejora del rendimiento ... 7224
Recopile y utilice estadísticas de tablas ... 7226
Desafíos habituales a la hora de escalar las cargas de trabajo de Trino ... 7227
Consideraciones sobre Trino ... 7232
Propiedades de la implementación de Trino no configurables ... 7232
Consideraciones adicionales ... 7233
Historial de lanzamientos de Trino ... 7233
Notas de la versión de Trino por versión ... 7233
Amazon EMR 7.6.0: notas de la versión de Trino ... 7234
Amazon EMR 7.3.0: notas de la versión de Trino ... 7234
Amazon EMR 6.9.0: notas de la versión de Trino ... 7234
Zeppelin ... 7236
Consideraciones sobre el uso de Zeppelin en Amazon EMR ... 7238

Historial de versiones de Zeppelin 7239
ZooKeeper 7301
ZooKeeper historial de versiones 7303
ZooKeeper notas de la versión 7344
Amazon EMR 7.8.0: notas de la versión ZooKeeper 7344
Amazon EMR 7.5.0: notas de la versión ZooKeeper 7344
Amazon EMR 7.4.0: notas de la versión ZooKeeper 7345
Conectores y utilidades 7346
Exportar, consultar y unir tablas en DynamoDB 7346
Configurar una tabla de Hive para ejecutar comandos de Hive 7348
Ejemplo de comandos de Hive para la exportación, importación y consulta de datos 7357
Optimización del rendimiento 7367
Kinesis 7371
¿Qué puede hacer con la integración de Amazon EMR y Amazon Kinesis? 7371
Análisis con puntos de control de flujos de Amazon Kinesis 7371
Consideraciones sobre el rendimiento 7373
Programación del análisis de Amazon Kinesis con Amazon EMR 7373
Migración del conector de Spark Kinesis al SDK 2.x para Amazon EMR 7.0 7373
S3 (s3-dist-cp) DistCp 7383
DistCp Opciones de S3 7384
Añadir S3 DistCp como un paso de un clúster 7392
Limpiando después de DistCp trabajos fallidos en S3 7394
Ejecución de comandos y scripts en un clúster 7396
Envío de un paso JAR personalizado para ejecutar un script o un comando 7397
Otras formas de usar `command-runner.jar` 7398
AWS Glosario 7400
..... 7401

Acerca de las versiones de Amazon EMR

Una versión de Amazon EMR es un conjunto de aplicaciones de código abierto del ecosistema de macrodatos. Cada versión incluye diferentes aplicaciones, componentes y características de macrodatos que selecciona para que Amazon EMR los instale y configure al crear un clúster. Las aplicaciones se empaquetan utilizando un sistema basado en Apache BigTop, que es un proyecto de código abierto asociado al ecosistema Hadoop. Esta guía proporciona información para las aplicaciones incluidas en las versiones de Amazon EMR.

Para obtener más información sobre cómo empezar a trabajar con Amazon EMR, consulte la Guía de administración de Amazon EMR.

Al lanzar un clúster, puede elegir entre varias versiones de Amazon EMR. Esto le permite probar y utilizar versiones de la aplicación que cumplan con sus requisitos de compatibilidad. El número de versión se especifica con la etiqueta de versión. Las etiquetas de versión tienen el formato `emr-x.x.x` Por ejemplo, `emr-7.9.0`.

A partir de Amazon EMR 5.18.0, puede utilizar el repositorio de artefactos de Amazon EMR para compilar el código de trabajos con las versiones exactas de las bibliotecas y dependencias que están disponibles con determinadas versiones de Amazon EMR. Para obtener más información, consulte Comprobación de dependencias mediante el repositorio de artefactos de Amazon EMR.

Para recibir actualizaciones cuando haya una nueva versión de Amazon EMR disponible, suscríbase a la fuente RSS de las notas de la versión de Amazon EMR.

Detalles de la última versión, incluidas las versiones de las aplicaciones, las notas de la versión, los componentes y las clasificaciones de configuración de las series 7.x, 6.x y 5.x de Amazon EMR:

- Amazon EMR versión 7.9.0
- Amazon EMR, versión 6.15.0
- Amazon EMR, versión 5.36.2

> **Note**
>
> Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento

inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Notas de la versión de las últimas versiones de Amazon EMR y un historial de todas las versiones:

- Novedades
- Archivo de notas de la versión de Amazon EMR

Un completo historial de versiones de la aplicación en cada versión de Amazon EMR:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Detalles de cada versión de Amazon EMR y diferencias entre las series de lanzamiento, si procede:

- Versiones de Amazon EMR 7.x
- Versiones de lanzamiento de Amazon EMR 6.x
- Versiones de lanzamiento de Amazon EMR 5.x
- Versiones de lanzamiento de Amazon EMR 4.x
- Versiones 2.x y 3.x de la AMI de Amazon EMR

Soporte estándar de Amazon EMR

Descripción de las versiones de Amazon EMR

La versión Amazon EMR es el vehículo que ofrece el software necesario para ejecutar sus aplicaciones de código abierto en EC2 plataformas EKS o sin servidor. Una versión de Amazon EMR se compone de tres partes: entorno de tiempo de ejecución, motores principales y extras.

- Entorno de tiempo de ejecución: el entorno de tiempo de ejecución incluye el sistema operativo que se incluye como parte de la Imagen de máquina de Amazon (AMI) u otra imagen de contenedor al lanzar el clúster o el contenedor (por ejemplo, Amazon Linux 2023). También incluye

los tiempos de ejecución del idioma, como el Amazon Corretto JDK, junto con otras herramientas relacionadas que instala la imagen.

- Motores principales: incluyen el software principal de código abierto, como Apache Hive y los formatos de tabla abierta, como Apache Iceberg.
- Extras: incluyen bibliotecas prácticas y paquetes de Python, así como `mariadb-connector-java` y software de código abierto como Apache Pig.

La versión de lanzamiento de Amazon EMR que seleccione agrupa todos estos componentes en una nueva versión de la Imagen de máquina de Amazon (AMI) o imágenes de contenedor basadas en Amazon Linux (AL). Amazon EMR mantiene la misma versión AL para todos los nodos de un clúster, un contenedor EKS o una aplicación sin servidor. El objetivo de Amazon EMR es disponer de la versión más reciente del entorno de tiempo de ejecución en un plazo de 90 días a partir de su lanzamiento por parte de Amazon Linux, así como de las nuevas versiones de código abierto de los motores principales en un plazo de 90 días a partir de su lanzamiento previo. Los extras se publican según sea necesario. Puede encontrar la lista del software en las notas de la versión en cada versión de Amazon EMR.

Control de versiones de lanzamiento

Para ayudarle a comprender el alcance de los cambios en cada versión, Amazon EMR utiliza el control de versiones semántico. El control de versiones semántico sigue un formato de `<major>.<minor>.<patch>`. Como ejemplo, consideremos la versión 7.0.0.

El primer dígito indica la versión principal publicada o versión principal. Las versiones principales suelen incluir cambios, mejoras y nuevas características sustanciales en entorno de tiempo de ejecución o motor principal que no son compatibles con versiones anteriores. Por ejemplo, la versión principal 7.x usa Amazon Linux 2023 con Amazon Corretto 17 JDK de forma predeterminada. Como resultado, la versión contiene varios cambios importantes en comparación con la versión principal 6.x, que se incluye en Amazon Linux 2 y con Amazon Corretto 8 JDK de forma predeterminada.

El segundo dígito indica la versión secundaria publicada o versión secundaria. Las versiones secundarias son versiones no compatibles con versiones anteriores que contienen cambios incrementales, mejoras y características en los motores principales y los extras, así como nuevas funcionalidades. Por ejemplo, la versión 6.15 usa Apache Spark 3.4.1, en comparación con la 6.11.1, que viene incluida con Apache Spark 3.3.2.

El tercer dígito indica la versión de parche publicada o versión de parche. Las versiones de parches están pensadas para ser versiones compatibles con versiones anteriores que contengan

correcciones y actualizaciones de parches para el motor principal y los extras, pero no nuevas funcionalidades ni actualizaciones de versiones del OSS. Por ejemplo, tanto la 6.11.1 como la 6.11.0 contienen la misma versión del OSS Apache Spark 3.3.2. A partir de la versión 6.6.x, la última versión del parche incluye el último parche del entorno de tiempo de ejecución disponible. El entorno de tiempo de ejecución no afecta a la versión del parche de Amazon EMR. Por ejemplo, la versión 6.11.1 se lanzó con la versión 2.0.20240109.0 de AL del 19 de diciembre de 2023 al 10 de enero de 2024, pero con la versión 2.0.20240124.0 de AL el 11 de enero de 2024. Algunas versiones de Amazon EMR en EKS le permiten añadir etiquetas o sufijos a las versiones semánticas para proporcionarle funcionalidades alternativas. Por ejemplo, la versión `emr-6.15.0-latest` de Amazon EMR en EKS lanza una versión de lanzamiento habilitada para Amazon Corretto 8 JDK, mientras que la versión `emr-6.15.0-java17-latest` de Amazon EMR en EKS permite lanzar una versión de lanzamiento habilitada para Amazon Corretto 17 JDK.

Política de soporte

Introducción

Amazon EMR tiene como objetivo publicar nuevas versiones secundarias como mínimo una vez cada 90 días y ofrecer soporte para las versiones secundarias durante un período de 24 meses a partir de la fecha de lanzamiento. Este soporte cubre el entorno de tiempo de ejecución y los motores principales con sus dependencias asociadas, y no cubre los extras ni sus dependencias asociadas. El entorno de ejecución también incluye componentes de plataforma patentados APIs que son necesarios para iniciar, detener y operar los clústeres y las aplicaciones. Este programa de lanzamiento coherente garantiza un ciclo predecible, lo que facilita la planificación, las pruebas y la transición a una versión que ofrezca soporte y seguridad.

Qué esperar del soporte estándar

El soporte estándar proporciona correcciones en los tickets de asistencia técnica para los problemas que surjan con los componentes del entorno de tiempo de ejecución y los motores principales en las configuraciones recomendadas. Todas las correcciones están sujetas a disponibilidad. La configuración recomendada es el uso de Amazon EMR sin ninguna modificación, adición o cambio en los binarios y configuraciones presentes en una versión de Amazon EMR, excepto los que se encuentren en la documentación de Amazon EMR. Amazon EMR implementa las correcciones en el último parche o en versión secundaria o principal de la versión de Amazon EMR en un plazo de 90 días a partir de la verificación de la corrección. Amazon EMR aplica correcciones automáticamente cuando lanza un nuevo EMR en un EC2 clúster, lanza un nuevo Amazon EMR en un contenedor

EKS o activa un nuevo trabajo EMR Serverless. Los componentes extras se proporcionan para su comodidad y Amazon EMR no proporciona correcciones relacionadas con los extras.

Componentes

El soporte estándar cubre los componentes de entorno de tiempo de ejecución y los motores principales, por ejemplo, el sistema operativo, los tiempos de ejecución de idiomas y el software principal de código abierto, como Apache Hive y Apache Iceberg. Puede encontrar la lista completa de componentes compatibles para cada versión en Acerca de las versiones de Amazon EMR Amazon Linux FAQs.

La siguiente lista describe el soporte que ofrecemos para los diferentes tipos de componentes dentro del soporte estándar:

- Componentes del entorno de tiempo de ejecución: los componentes del entorno de tiempo de ejecución recibirán correcciones en los tickets de asistencia técnica. Las correcciones se clasifican en (a) errores críticos, (b) problemas críticos de corrupción de datos y (c) problemas críticos de seguridad. Cuando sea elegible, Amazon EMR realizará una portabilidad con correcciones para el entorno de tiempo de ejecución a versiones anteriores. Para garantizar la compatibilidad operativa con los componentes de código abierto del motor principal, algunos componentes del entorno de tiempo de ejecución deben permanecer en versiones específicas para evitar problemas en las aplicaciones de los clientes. Para estos componentes, Amazon EMR depende del código abierto original para la disponibilidad de las correcciones. Cuando las correcciones estén disponibles en código abierto, proporcionaremos la última versión estable en un plazo de 90 días a partir de su verificación por parte de Amazon EMR.
- Componentes de los motores principales: los motores principales proporcionan las versiones más recientes para muchos proyectos de código abierto, cada uno con cientos de bibliotecas de dependencias transitivas. Si bien las comunidades de código abierto que gestionan estos proyectos intentan abordar los problemas y las vulnerabilidades y exposiciones comunes (CVE) conocidas con frecuencia, es posible que las versiones más recientes aún contengan errores conocidos y. CVEs Amazon EMR depende del código abierto original para la disponibilidad de las correcciones y proporcionará la última versión estable como parte de los componentes del motor principal en un plazo de 90 días a partir de su verificación por parte de Amazon EMR. En algunos casos, Amazon EMR puede proporcionar una corrección para un CVE en uno de los componentes del motor principal que debe abordarse con anterioridad al código abierto previo. Amazon EMR también le proporciona asistencia técnica y correcciones en las características que se añaden al código abierto según las configuraciones recomendadas. No realizamos

portabilidades de correcciones de los componentes del motor principal a versiones anteriores de parches o versiones secundarias.

- Componentes extras: Amazon EMR no admite componentes extras. Los componentes extras son proyectos de código abierto que se proporcionan para su comodidad y Amazon EMR no proporciona correcciones para los problemas encontrados con ellos. Cualquier solicitud de soporte o corrección puede abordarse a través de la comunidad de código abierto que respalda estos componentes.

Ciclo de vida del soporte estándar

A continuación se describen los hitos del ciclo de vida del soporte estándar:

- Soporte estándar: las versiones de Amazon EMR son aptas para el soporte estándar 24 meses después de la fecha de lanzamiento. Puede crear tickets de asistencia técnica y esperar actualizaciones para los problemas que surjan con estas versiones.
- Fin del soporte: una vez finalizado el soporte estándar, las versiones de Amazon EMR entran en la fase de fin de soporte (EoS) durante 12 meses. Las versiones de EoS no reúnen los requisitos para recibir asistencia técnica y no podrá crear ningún ticket para clústeres, contenedores o trabajos que se ejecuten en estas versiones. Las versiones de EoS no recibirán correcciones, parches ni actualizaciones. Las versiones de EoS se eliminarán de la consola, pero seguirán estando disponibles a través de la API y la AWS CLI. Siempre podrá seguir ejecutando cargas de trabajo en las versiones de EoS. Recomendamos encarecidamente migrar a la versión más reciente de Amazon EMR para que continúe recibiendo parches de seguridad, siga siendo elegible para asistencia técnica y pueda crear tickets de soporte cuando sea necesario.
- Fin de vida útil: después del período de fin de soporte (EoS), las versiones se consideran en la fase de fin de vida útil (EoL). Si bien puede seguir ejecutando clústeres de EoL, Amazon EMR se reserva el derecho de eliminar las versiones de EoL de la API y el case-by-case SDK por motivos operativos y de seguridad. Le recomendamos encarecidamente que migre a la versión más reciente de Amazon EMR, ya que las versiones de EoL se pueden eliminar de la API y el SDK en casos excepcionales.

Bridge Support

Amazon EMR anunció esta nueva política de soporte el 25 de julio de 2024. Según esta política, las versiones de Amazon EMR publicadas el 24 de julio de 2022 o antes ahora estarán en la fase de fin de soporte. Sin embargo, para proporcionarle más tiempo para planificar y migrar a versiones

más recientes, Amazon EMR ofrecerá Bridge Support, equivalente al soporte estándar, para estas versiones anteriores que se publiquen en el período de dos años previo a este anuncio.

A partir del 25 de julio de 2024, puede consultar el estado actual y los plazos de soporte en las notas de la versión.

En la siguiente tabla se muestra el estado de soporte de todas las versiones actuales de Amazon EMR en el momento del anuncio de la política, el 25 de julio de 2024:

Versiones y períodos con soporte

Versiones de lanzamiento de Amazon EMR	Fecha de lanzamiento inicial	Fecha de fin del soporte estándar	Fecha de inicio del fin de soporte	Fecha de inicio del fin de vida útil
7.2.0	25 de julio de 2024	24 de julio de 2026	25 de julio de 2026	25 de julio de 2027
7.1.0	23 de abril de 2024	22 de abril de 2026	23 de abril de 2026	23 de abril de 2027
7.0.0	19 de diciembre de 2023	18 de diciembre de 2025	19 de diciembre de 2025	19 de diciembre de 2026
5.36.x y 6.6.x-6.15.x	Del 9 de mayo de 2022 al 13 de noviembre de 2023	Bridge Support hasta el 24 de enero de 2026	25 de enero de 2026	25 de enero de 2026
• Serie 6.x: 6.5.0 y versiones anteriores • Serie 5.x: 5.35.0 y versiones anteriores • Series 4.x, 3.x y 2.x	Del 1 de enero de 2013 al 30 de marzo de 2022	Bridge Support hasta el 24 de julio de 2025	25 de julio de 2025	25 de julio de 2025

Consideraciones

El soporte estándar está disponible para todos los modelos de implementación de Amazon EMR (EMR activado, EC2 Amazon EMR en EKS y EMR Serverless), en todas las regiones en las que Amazon EMR esté disponible, sin coste adicional. Los clústeres que se ejecutan con las configuraciones recomendadas pueden recibir soporte automáticamente, tal como se describe en la política, por lo que no es necesario realizar ninguna acción adicional para activar el soporte.

- El soporte estándar solo admite componentes necesarios para clústeres de Amazon EMR. Amazon EMR no puede garantizar la disponibilidad de parches y correcciones de seguridad en el caso de que los componentes de código abierto del motor principal lleguen al EoL original o cuando las actualizaciones de seguridad ya no estén disponibles para las dependencias. Si bien puede optar por instalar extras, Amazon EMR no admitirá estos ni tampoco sus dependencias. Por ejemplo, puede instalar aplicaciones de terceros en su AMI personalizada para reforzar la seguridad de su clúster, instalar componentes adicionales o copiar objetos mediante scripts de acción de arranque o SSH en su clúster y actualizar las versiones predeterminadas del paquete. Amazon EMR no admite estos componentes. El soporte estándar no cubre las acciones de bootstrap, los paquetes, las bibliotecas ni el código personalizado ni las aplicaciones bring-your-own personalizadas que Amazon EMR puede configurar para que se instalen según su comodidad.
- Sus clústeres existentes no se verán afectados, independientemente de la versión de Amazon EMR que estén ejecutando. Puede seguir ejecutando los clústeres existentes sin interrupciones. También puede seguir lanzando nuevos clústeres y ejecutando trabajos en cualquiera de las versiones existentes y en las nuevas versiones. Todas las versiones existentes y las nuevas versiones en el momento en que la política entre en vigor están cubiertas por el soporte estándar durante 24 meses a partir de la fecha inicial de publicación de la versión de lanzamiento de Amazon EMR. Amazon EMR le proporcionará Bridge Support durante el anuncio inicial de la política. Para recibir soporte ininterrumpido, le recomendamos que pruebe sus aplicaciones y las actualice rápidamente a la versión más reciente de Amazon EMR.
- Amazon EMR no cambiará los componentes del soporte estándar en las versiones o clústeres existentes. Sin embargo, Amazon EMR se reserva el derecho de respetar el fin del ciclo de vida anterior y eliminar dichos componentes en las nuevas versiones, o pasar la versión existente al estado de End of Support (EoS) o End of Life (EoL) en casos excepcionales. case-by-case Le notificaremos sobre cualquier eliminación a través de los canales disponibles.

- Cuando se lance una nueva región, Amazon EMR solo admitirá las versiones de lanzamiento de Amazon EMR con soporte estándar, publicadas seis meses antes de la fecha en que la nueva región esté disponible para el público en general.
- Amazon EMR no actualizará automáticamente los clústeres existentes a las versiones más recientes. Sin embargo, si lo desea, puede optar por actualizar los nuevos clústeres a las últimas versiones del parche.
- El coste del uso no admitido de Amazon EMR se tendrá en cuenta en la factura. AWS Incluso si utiliza Amazon EMR de forma no admitida, los costes asociados a ese uso seguirán formando parte de su AWS consumo total y se incluirán en el cálculo de las tarifas de soporte.

Para obtener más información, póngase en contacto con AWS Developer Support.

Versiones de Amazon EMR 7.x

Esta sección contiene las versiones de las aplicaciones, las notas de la versión, las versiones de los componentes y las clasificaciones de configuración de cada versión 7.x de Amazon EMR.

Al lanzar un clúster, puede elegir entre varias versiones de Amazon EMR. Esto le permite probar y utilizar versiones de la aplicación que cumplan con sus requisitos de compatibilidad. El número de versión se especifica con la etiqueta de versión. Las etiquetas de versión tienen el formato `emr-x.x.x` Por ejemplo, `emr-7.9.0`.

Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Para ver una tabla completa de versiones de las aplicaciones en cada versión 7.x de Amazon EMR, consulte Versiones de las aplicaciones en las versiones 7.x de Amazon EMR.

Temas

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Amazon EMR versión 7.9.0
- Amazon EMR versión 7.8.0
- Amazon EMR versión 7.7.0
- Amazon EMR versión 7.6.0

- Amazon EMR versión 7.5.0
- Amazon EMR versión 7.4.0
- Amazon EMR, versión 7.3.0
- Amazon EMR, versión 7.2.0
- Versión 7.1.0 de Amazon EMR
- Amazon EMR, versión 7.0.0

Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

Para ver una tabla completa con las versiones de las aplicaciones disponibles en cada versión 7.x de Amazon EMR, abra Versiones de las aplicaciones en las versiones 7.x de Amazon EMR en su navegador.

Amazon EMR versión 7.9.0

Ciclo de vida compatible con 7.9.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.9.0 de Amazon EMR.

Fase de soporte	Date
Fecha de lanzamiento inicial	19 de mayo de 2025
Soporte estándar hasta	19 de mayo de 2027
Fin del soporte	20 de mayo de 2027
Fin de vida útil	19 de mayo de 2028

componentes de la versión 7.9.0

Amazon EMR 7.9.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

Hue, Oozie, Pig, Presto, Zeppelin JupyterHub TensorFlow

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de la aplicación 7.9.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.9.0	emr-7.8.0	emr-7.7.0	emr-7.6.0
AWS SDK for Java	2.31.16, 1.12.782	2.29.52, 1.12.780	2.29.44, 1.12.780	2.29.25, 1.12.779
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300032,2

	emr-7.9.0	emr-7.8.0	emr-7.7.0	emr-7.6.0
Delta	3.3.0-amzn-1	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1
Flink	1.20.0	1.20.0	1.20.0	1.20.0
HBase	2.6.2	2.6.1	2.6.1	2.6.1
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.1	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-6	0.15.0-amzn-5	0.15.0-amzn-4	0.15.0-amzn-3
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-2	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.1	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.5	3.5.4	3.5.3	3.5.3
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-7.9.0	emr-7.8.0	emr-7.7.0	emr-7.6.0
Trino (PrestoSQL)	467	467	467	457
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.3	3.9.2	3.9.2

notas de la versión 7.9.0

Las siguientes notas de la versión incluyen información sobre la versión 7.9.0 de Amazon EMR.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.
- Compatibilidad con CM-CMK para la interfaz de usuario persistente de aplicaciones: esta función amplía el cifrado SSE aplicado a los registros gestionados en el depósito de servicios de EMR para cumplir con la clave maestra personalizada gestionada por el cliente (CM-CMK). Esta función permite a los clientes gestionar la accesibilidad de los registros una vez finalizados sus clústeres.

Problemas conocidos

- Cuando la opción de registro para el EMR activado EC2 está deshabilitada, los registros no están disponibles una vez finalizado el clúster y la aplicación persistente no UIs contendrá ningún registro.

Cambios, mejoras y problemas resueltos

- Obsolescencia de Oozie: dado que OSS Oozie ha pasado al modo ático, la última versión que incluye Apache Oozie es la versión 7.9 de Amazon EMR. Apache Oozie se eliminará de la versión 7.10 y versiones posteriores de Amazon EMR. Le recomendamos que utilice MWAA como alternativa a Oozie. Los clientes pueden utilizar la utilidad para migrar los flujos oozie-to-airflow-emrde trabajo de Oozie existentes GitHub a MWAA.

- Embalaje de contenedores YARN: a partir de la versión 7.9.0 de Amazon EMR, la política de embalaje de contenedores ya está disponible para el programador de capacidad de YARN, que se basa en la política de ubicación de varios nodos de YARN.
- El archivo hudi-cli.sh: el script hudi-cli.sh quedó obsoleto en la versión 7.9.0 y posteriores de EMR. La versión 7.9.0 y superior de Amazon EMR usa .jar. hudi-cli-bundle
- Clústeres de EMR en una AMI AL2 023 habilitada para FIPS: a partir de EMR 7.9, los clientes pueden lanzar clústeres de EMR en una AMI 023 habilitada para FIPS. AL2 Este cambio actualiza las firmas digitales y los algoritmos de hash de EMR para garantizar que todos los paquetes de EMR cumplan con las políticas criptográficas recomendadas para el FIPS 140-3. Nota: Esto no cubre el estado de la certificación FIPS de EMR AL2 o 023; consulte las preguntas frecuentes sobre la FIPS 023 AL2para obtener la información más reciente sobre el estado de validación de la FIPS 140-3.
- Latencia escalonada: esta versión resuelve un problema por el que la latencia escalonada aumentaba cuando se utilizaba el cifrado en tránsito sin seleccionar explícitamente la aplicación Hadoop para su instalación.
- Lanzamiento de clústeres con volúmenes cifrados con EBS: esta versión corrige un problema por el que una llamada de KMS en volúmenes cifrados con EBS podía provocar un error en el lanzamiento de un clúster de EMR.

-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyu Derabad), EE.UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

Versiones Java predeterminadas de 7.9.0

La versión 7.9 y posteriores de Amazon EMR incluyen Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla se muestran las versiones predeterminadas de Java para las aplicaciones de Amazon EMR 7.9.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
ZooKeeper	17, 11, 8

Versiones de componentes 7.9.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de

macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1.31.0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.3.0-amzn-1	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.18.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.22.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.18.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.39.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.18.0	Conector S3Select de EMR
emr-wal-cli	1.3.4	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.70,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.20.0-amzn-3	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.20.0-amzn-3	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.4.1-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.4.1-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.4.1-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.4.1-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.4.1-amzn-1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.4.1-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.4.1-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.4.1-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.4.1-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.4.1-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.4.1-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.6.2-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.6.2-amzn-0	Servicio para prestar servicio a una o más regiones. HBase
hbase-client	2.6.2-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.6.2-amzn-0	Servicio que proporciona un punto final HTTP para. RESTful HBase
hbase-thrift-server	2.6.2-amzn-0	Servicio que proporciona un punto final de Thrift para. HBase
hbase-operator-tools	2.6.2-amzn-0	Herramienta de reparación de clústeres de Apache. HBase
hcatalog-client	3.1.3-amzn-18	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-18	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-18	Punto final HTTP que proporciona una interfaz REST para. HCatalog
hive-client	3.1.3-amzn-18	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-18	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-18	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-18	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.15.0-amzn-6	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.15.0-amzn-6	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.15.0-amzn-6	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.15.0-amzn-6	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.7.1-amzn-2	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	12.3.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.2.1	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.287-amzn-3	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.287-amzn-3	Service para ejecutar partes de una consulta.
presto-client	0.287-amzn-3	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	467-amzn-2	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	467-amzn-2	Service para ejecutar partes de una consulta.
trino-client	467-amzn-2	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.5-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.5.5-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.5-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.5-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	25.02.1-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-16	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
tez-on-worker	0.10.2-amzn-16	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.11.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.3-amzn-1	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.3-amzn-1	ZooKeeper cliente de línea de comandos.

clasificaciones de configuración de 7.9.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

clasificaciones emr-7.9.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.

Clasificaciones	Descripción	Acciones de reconfiguración
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j2	Cambie los valores en el archivo HBase hbase-log 4j2.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j2	Cambie los valores en el archivo log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml del Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.

Clasificaciones	Descripción	Acciones de reconfiguración
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

7.9.0 registro de cambios

Registro de cambios de la versión 7.9.0 y de las notas de la versión

Date	Evento	Descripción
2025-05-19	Publicación de documentos	Las notas de la versión 7.9.0 de Amazon EMR se publicaro n por primera vez
2025-05-19	Implementación completa	Amazon EMR 7.9.0 se ha implementado completam ente en todas las regiones compatibles
2025-05-19	Versión limitada inicial	Amazon EMR 7.9.0 se implementó por primera vez en las regiones comerciales iniciales

Amazon EMR versión 7.8.0

Ciclo de vida compatible con 7.8.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.8.0 de Amazon EMR.

Fase de soporte	Date	
Fecha de lanzamiento inicial	7 de marzo de 2025	
Soporte estándar hasta	6 de marzo de 2027	

Fase de soporte	Date
Fin del soporte	7 de marzo de 2027
Fin de vida útil	7 de marzo de 2028

componentes de la versión 7.8.0

Amazon EMR 7.8.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

Hue, Oozie, Pig, Presto, Zeppelin JupyterHub TensorFlow

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de la aplicación 7.8.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.8.0	emr-7.7.0	emr-7.6.0	emr-7.5.0
AWS SDK for Java	2.29.52, 1.12.780	2.29.44, 1.12.780	2.29.25, 1.12.779	2.28,8, 1.12.772
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0
Flink	1.20.0	1.20.0	1.20.0	1.19.1
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.6.1	2,5,10
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-5	0.15.0-amzn-4	0.15.0-amzn-3	0.15.0-amzn-2
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-

	emr-7.8.0	emr-7.7.0	emr-7.6.0	emr-7.5.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.4	3.5.3	3.5.3	3.5.2
Sqoop	-	-	-	-
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	467	457	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.2	3.9.2	3.9.2

notas de la versión 7.8.0

Las siguientes notas de la versión incluyen información sobre la versión 7.8.0 de Amazon EMR.

Nuevas características

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

Problemas conocidos

- No se aplica a esta versión.

Cambios, mejoras y problemas resueltos

- No hay cambios importantes.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyu Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,1	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

Versiones Java predeterminadas de 7.8.0

La versión 7.8 y las versiones posteriores de Amazon EMR incluyen Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla se muestran las versiones predeterminadas de Java para las aplicaciones de Amazon EMR 7.8.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
ZooKeeper	17, 11, 8

Versiones de componentes 7.8.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de

macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1.31.0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.3.0-amzn-0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.17.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.21.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.18.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.38.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.17.0	Conector S3Select de EMR
emr-wal-cli	1.3.3	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.69,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.20.0-amzn-2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.20.0-amzn-2	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.4.1-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.4.1-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.4.1-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.4.1-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.4.1-amzn-0	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.4.1-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.4.1-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.4.1-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.4.1-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.4.1-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.4.1-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.6.1-amzn-2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.6.1-amzn-2	Servicio para prestar servicio a una o más regiones. HBase
hbase-client	2.6.1-amzn-2	HBase cliente de línea de comandos.
hbase-rest-server	2.6.1-amzn-2	Servicio que proporciona un punto final HTTP para. RESTful HBase
hbase-thrift-server	2.6.1-amzn-2	Servicio que proporciona un punto final de Thrift para. HBase
hbase-operator-tools	2.6.1-amzn-2	Herramienta de reparación de clústeres de Apache. HBase
hcatalog-client	3.1.3-amzn-17	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-17	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-17	Punto final HTTP que proporciona una interfaz REST para. HCatalog
hive-client	3.1.3-amzn-17	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-17	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-17	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-17	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.15.0-amzn-5	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.15.0-amzn-5	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.15.0-amzn-5	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.15.0-amzn-5	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.7.1-amzn-1	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	12.3.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.2.1	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.287-amzn-2	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.287-amzn-2	Service para ejecutar partes de una consulta.
presto-client	0.287-amzn-2	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	467-amzn-1	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	467-amzn-1	Service para ejecutar partes de una consulta.
trino-client	467-amzn-1	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.4-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.5.4-amzn-0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.4-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.4-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	24.12.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-15	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
tez-on-worker	0.10.2-amzn-15	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.11.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.3-amzn-0	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.3-amzn-0	ZooKeeper cliente de línea de comandos.

clasificaciones de configuración de 7.8.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

clasificaciones emr-7.8.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.

Clasificaciones	Descripción	Acciones de reconfiguración
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j2	Cambie los valores en el archivo log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml del Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.

Clasificaciones	Descripción	Acciones de reconfiguración
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

7.8.0 registro de cambios

Registro de cambios de la versión 7.8.0 y de las notas de la versión

Date	Evento	Descripción
07-03-2025	Publicación de documentos	Las notas de la versión 7.8.0 de Amazon EMR se publicaro n por primera vez
06/03/2025	Implementación completa	Amazon EMR 7.8.0 se ha implementado completam ente en todas las regiones compatibles
28-02-2021	Versión limitada inicial	Amazon EMR 7.8.0 se implementó por primera vez en las regiones comerciales iniciales

Amazon EMR versión 7.7.0

Ciclo de vida compatible con 7.7.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.7.0 de Amazon EMR.

Fase de soporte	Date	
Fecha de lanzamiento inicial	6 de febrero de 2025	
Soporte estándar hasta	5 de febrero de 2027	

Fase de soporte	Date
Fin del soporte	6 de febrero de 2027
Fin de vida útil	6 de febrero de 2028

componentes de la versión 7.7.0

Amazon EMR 7.7.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

Hue, Oozie, Pig, Presto, Zeppelin JupyterHub TensorFlow

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de la aplicación 7.7.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
AWS SDK for Java	2.29.44, 1.12.780	2.29.25, 1.12.779	2.28,8, 1.12.772	2.28.8, 1.12.772
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1
Flink	1.20.0	1.20.0	1.19.1	1.19.1
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.5.10	2.5.5
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-4	0.15.0-amzn-3	0.15.0-amzn-2	0.15.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-

	emr-7.7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.0	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,287
Spark	3.5.3	3.5.3	3.5.2	3.5.2
Sqoop	-	-	-	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	457	446	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.2

Detalles de los componentes de la versión 7.7.0

Consulte la siguiente tabla para obtener más información sobre los paquetes Extras de Amazon EMR 7.7.0.

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	Extras	3.4.0-amzn-3-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	Extras	2.0.0-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	s3-dist-cp	Extras	2.37.0-1.amzn2023
AmazonCloudWatchAgent	Core	ZooKeeper	Extras	3.9.2. amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	Extras	3.9.2. amzn.1-1.amzn2023
Flink	Core	hadoop-client	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-httpfs	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-kms	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Flink	Core	hadoop-yarn	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.3-1.amzn2023
Flink	Core	hudi	Extras	0.15.0.amzn.4-1.amzn2023
Flink	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Core	hadoop	Extras	3.4.0.amzn.3-1.amzn2023
HBase	Core	s3-dist-cp	Extras	2.37.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	Extras	3.4.0. amzn.3-1.amzn2023
HCatalog	Core	mariadb-connector-java	Extras	

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Hive	Core	hadoop-httpfs	Extras	3.4.0.amzn.3-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Hive	Core	hive-hbase	Extras	3.1.3. amzn.16-1.amzn2023
Hive	Core	mariadb-connector-java	Extras	2.7.2-1
Hive	Core	s3-dist-cp	Extras	2.37.0-1.amzn2023
Phoenix	Core	s3-dist-cp	Extras	2.37.0-1.amzn2023
Spark	Core	spark-datanucleus	Extras	3.5.3. amzn.1-1.amzn2023
Tez	Core	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
Tez	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023
Tez	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Core	hadoop-client	Extras	3.4.0.amzn.3-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-hdfs	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-kms	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.3-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hive	Extras	3.1.3. amzn.16-1.amzn2023
Trino	Core	hive-hcatalog-server	Extras	3.1.3. amzn.16-1.amzn2023
Trino	Core	mariadb-connector-java	Extras	2.7.2-1
Trino	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Core	ZooKeeper	Extras	3.9.2. amzn.1-1. amzn2023
Trino	Core	zookeeper-server	Extras	3.9.2. amzn.1-1. amzn2023
Trino	Core	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Core	hadoop	Extras	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-lzo	Extras	0.4.19-1. amzn2023
Trino	Core	hive-hcatalog	Extras	3.1.3. amzn.16-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hive-jdbc	Extras	3.1.3. amzn.16-1.amzn2023
ZooKeeper	Core	emrfs	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-client	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-hdfs	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-datanode	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-journalnode	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-namenode	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-httpfs	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-kms	Extras	3.9.2. amzn.1-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce	Extras	3.4.0.amzn.3-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.3-1.amzn2023
ZooKeeper	Core	hadoop-yarn	Extras	3.4.0.amzn.3-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
ZooKeeper	Core	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.3-1.amzn2023
ZooKeeper	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.3-1.amzn2023
ZooKeeper	Core	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.3-1.amzn2023
ZooKeeper	Core	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.3-1.amzn2023
ZooKeeper	Core	ranger-kms	Extras	2.0.0-1.amzn2023

Notas de la versión 7.7.0

Las siguientes notas de la versión incluyen información sobre la versión 7.7.0 de Amazon EMR.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.7.0 application upgrades include Delta 3.2.1-amzn-2, Hudi 0.15.0-amzn-4, Iceberg 1.7.1-amzn-0, and Trino 467.

Problemas conocidos

- No se aplica a esta versión.

Cambios, mejoras y problemas resueltos

- La versión de Iceberg que se utilizaba a partir de EMR 7.7.0 ya no es compatible con Java 8. Además, Iceberg está excluido de las siguientes imágenes de Java 8: y. `emr-7.7.0-java8-latest` `emr-7.7.0-spark-rapids-java8-latest`

- Lake Formation para FGAC con Amazon EMR en EKS: con la versión 7.7 y versiones posteriores de Amazon EMR, puede aprovechar AWS Lake Formation para aplicar controles de acceso detallados en las tablas del catálogo de datos respaldadas por Amazon S3. Esta capacidad le permite configurar controles de acceso a nivel de tabla, fila, columna y celda para las consultas de lectura en Amazon EMR de Amazon en EKS Spark Jobs. Para obtener más información, consulte Uso de Amazon EMR en EKS con AWS Lake Formation para obtener un control de acceso detallado.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyu Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,1	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

Versiones Java predeterminadas de 7.7.0

La versión 7.7 y posteriores de Amazon EMR incluyen Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla se muestran las versiones predeterminadas de Java para las aplicaciones de Amazon EMR 7.7.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
ZooKeeper	17, 11, 8

Versiones de componentes 7.7.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de

macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1.31.0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.2.1-amzn-2	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.16.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.20.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.18.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.37.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.16.0	Conector S3Select de EMR
emr-wal-cli	1.3.2	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.68,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.20.0-amzn-1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.20.0-amzn-1	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.4.0-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.4.0-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.4.0-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.4.0-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.4.0-amzn-3	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.4.0-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.4.0-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.4.0-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.4.0-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.4.0-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.6.1-amzn-1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.6.1-amzn-1	Servicio para prestar servicio a una o más regiones. HBase
hbase-client	2.6.1-amzn-1	HBase cliente de línea de comandos.
hbase-rest-server	2.6.1-amzn-1	Servicio que proporciona un punto final HTTP para. RESTful HBase
hbase-thrift-server	2.6.1-amzn-1	Servicio que proporciona un punto final de Thrift a. HBase
hbase-operator-tools	2.6.1-amzn-1	Herramienta de reparación de clústeres de Apache. HBase
hcatalog-client	3.1.3-amzn-16	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-16	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-16	Punto final HTTP que proporciona una interfaz REST para. HCatalog
hive-client	3.1.3-amzn-16	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-16	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-16	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-16	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.15.0-amzn-4	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.15.0-amzn-4	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.15.0-amzn-4	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.15.0-amzn-4	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.7.1-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	12.3.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.2.1	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.287-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.287-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.287-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	467-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	467-amzn-0	Service para ejecutar partes de una consulta.
trino-client	467-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.3-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	3.5.3-amzn-1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.3-amzn-1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.3-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	24.10.1-amzn-1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-14	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
tez-on-worker	0.10.2-amzn-14	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.11.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.2-amzn-1	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.2-amzn-1	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de 7.7.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

clasificaciones emr-7.7.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.

Clasificaciones	Descripción	Acciones de reconfiguración
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j2	Cambie los valores en el archivo log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml del Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.

Clasificaciones	Descripción	Acciones de reconfiguración
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

7.7.0 registro de cambios

Registro de cambios de la versión 7.7.0 y de las notas de la versión

Date	Evento	Descripción
2025-02-07	Publicación de documentos	Las notas de la versión 7.7.0 de Amazon EMR se publicaro n por primera vez
2025-02-06	Implementación completa	Amazon EMR 7.7.0 se ha implementado completam ente en todas las regiones compatibles
30-01-2020	Versión limitada inicial	Amazon EMR 7.7.0 se implementó por primera vez en las regiones comerciales iniciales

Amazon EMR versión 7.6.0

Ciclo de vida compatible con 7.6.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.6.0 de Amazon EMR.

Fase de soporte	Date	
Fecha de lanzamiento inicial	10 de enero de 2025	
Soporte estándar hasta	9 de enero de 2027	

Fase de soporte	Date
Fin del soporte	10 de enero de 2027
Fin de vida útil	10 de enero de 2028

componentes de la versión 7.6.0

Amazon EMR 7.6.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

Hue, Oozie, Pig, Presto, Zeppelin JupyterHub TensorFlow

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de la aplicación 7.6.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
AWS SDK for Java	2.29.25, 1.12.779	2.28,8, 1.12.772	2.28.8, 1.12.772	2.25,70, 1.12.747
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0
Flink	1.20.0	1.19.1	1.19.1	1.18.1
Ganglia	-	-	-	-
HBase	2.6.1	2.5.10	2.5.5	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-3	0.15.0-amzn-2	0.15.0-amzn-1	0.15.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.2.0	5.2.0	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,287	0,285
Spark	3.5.3	3.5.2	3.5.2	3.5.1
Sqoop	-	-	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	457	446	446	442
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.1

7.6.0 vulnerabilidades y exposiciones comunes

La siguiente tabla muestra todo lo CVEs que no afecta a los clústeres de EMR que se ejecutan en las configuraciones recomendadas de Amazon EMR 7.6.0. Amazon EMR depende del código abierto original para la disponibilidad de las correcciones y proporcionará la última versión estable como parte de los componentes del motor principal en un plazo de 90 días a partir de su verificación por parte de Amazon EMR.

Esta tabla se actualizó el 10 de enero de 2025.

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-15250	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-17521	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-17521
CVE-2024-12798	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-12798
CVE-2024-12801	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2024-12801
CVE-2024-47554	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-47554
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857
CVE-202238749	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238749
CVE-202238750	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238750
CVE-202238751	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238751
CVE-202238752	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238752
CVE-2022-41854	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-41854

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-36114	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-36114
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2018-1313	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-1313
CVE-2019-10219	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2019-10219
CVE-2020-10693	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-10693
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-13955	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-13955
CVE-2020-29582	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-29582
CVE-2020-8908	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2020-8908

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-28170	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-28170
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-38153	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-38153
CVE-202224329	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202224329
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-36364	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202236364
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202239135
CVE-2022-42889	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202242889
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2022-46751	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202246751
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-20863	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2023-20883	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-20883

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-25194	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-33201	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-33202	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-33202
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-42503	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2023-51074	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-51074

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2024-23945	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-23945
CVE-2024-25710	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2024-26308	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2024-29857	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-29857
CVE-2024-30171	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-30171
CVE-2024-31141	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-31141
CVE-2024-38808	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2024-38820	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-38820
CVE-2024-47561	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-47561
CVE-2024-52046	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-52046
CVE-2024-53990	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-53990
CVE-2024-56128	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-56128
CVE-2024-7254	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-7254

ID de CVE	Gravedad	URL de detalles del CVE
GHSA-CHFM-68VV-PVW5	BAJA	No hay ningún enlace disponible.
CVE-2024-47535	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-47535
CVE-2012-5783	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2012-5783
CVE-2012-6153	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2012-6153
CVE-2016-5018	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2018-10237	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-10237
CVE-2018-11798	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-11798
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-15522	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-15522
CVE-2020-1945	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2021-36373	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2021-36374	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36374

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-1000180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2020-26939	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2020-28052	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-29425	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-29425
CVE-2015-5237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2017-7525	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14719

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-12384	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2019-12384
CVE-2019-12814	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2019-12814
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-22569	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-22569
CVE-2021-22570	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2023-50387	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2023-50868	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-2023-52428	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-52428

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2024-23454	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2024-23944	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-23944
CVE-2024-25638	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-25638
CVE-2024-29025	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-29025
CVE-2024-29131	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-29131
CVE-2024-29133	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-29133
CVE-2024-51504	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2024-6763	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2024-8184	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2013-7285	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2013-7285
CVE-2014-8125	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2014-8125
CVE-2016-3674	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-3674
CVE-2017-7957	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-7957
CVE-2019-10173	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-10173
CVE-2020-26217	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-26217
CVE-2020-26258	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-26258
CVE-2021-21341	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-21341
CVE-2021-21342	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21342
CVE-2021-21343	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-21343
CVE-2021-21344	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21344
CVE-2021-21345	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21345

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-21346	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21346
CVE-2021-21347	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21347
CVE-2021-21348	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-21348
CVE-2021-21349	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-21349
CVE-2021-21350	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21350
CVE-2021-21351	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-21351
CVE-2021-29505	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-29505
CVE-2021-39139	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39139
CVE-2021-39141	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39141
CVE-2021-39144	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39144
CVE-2021-39145	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39145
CVE-2021-39146	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39146
CVE-2021-39147	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39147

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-39148	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39148
CVE-2021-39149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39149
CVE-2021-39150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39150
CVE-2021-39151	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39151
CVE-2021-39152	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39152
CVE-2021-39153	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39153
CVE-2021-39154	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-39154
CVE-2021-41411	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-41411
CVE-2021-43859	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-43859
CVE-2022-1415	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-1415
CVE-202240151	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240151
CVE-2022-41966	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-41966
CVE-2024-47072	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-47072

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-10172	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-10172
CVE-2019-10202	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-10202
CVE-2022-45688	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245688
CVE-2023-0833	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2024-21634	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2024-9823	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-9823
CVE-2022-37866	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202237866
CVE-2012-0881	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-202223437	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202223437
CVE-2020-13936	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2022-37865	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202237865
CVE-2009-2625	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2009-2625
CVE-2017-10355	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2017-10355

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2020-14338	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-14338
CVE-2018-2799	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-2799
CVE-2021-27568	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2013-4002	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2020-13956	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-13956
CVE-2020-25649	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2023-26049	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2024-43382	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2024-43382
CVE-2023-34462	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-34462
CVE-2021-21290	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2023-44487	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-44487
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20445

ID de CVE	Gravedad	URL de detalles del CVE
GHSA-XPW8-RCWV-8F8p	ALTO	No hay ningún enlace disponible.
CVE-2021-37136	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-21295	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-21295
GHSA-58QW-P7QM-5RVh	BAJA	No hay enlace disponible.
CVE-2021-37533	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-37533
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2022-40152	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240152

Detalles de los componentes de la versión 7.6.0

Consulte la siguiente tabla para obtener más información sobre los paquetes Extras de Amazon EMR 7.6.0.

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.2-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	Extras	2.36.0-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	ZooKeeper	Extras	3.9.2. amzn.0-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	Extras	3.9.2. amzn.0-1.amzn2023
Flink	Core	hadoop-client	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-httpfs	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-kms	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-mapreduce	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn	Extras	3.4.0.amzn.2-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Flink	Core	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.2-1.amzn2023
Flink	Core	hudi	Extras	0.15.0.amzn.3-1.amzn2023
Flink	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Core	hadoop	Extras	3.4.0.amzn.2-1.amzn2023
HBase	Core	s3-dist-cp	Extras	2.36.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	Extras	3.4.0. amzn.2-1.amzn2023
HCatalog	Core	mariadb-connector-java	Extras	
Hive	Core	hadoop-httpfs	Extras	3.4.0.amzn.2-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Hive	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
Hive	Core	hive-hbase	Extras	3.1.3. amzn.15-1.amzn2023
Hive	Core	mariadb-connector-java	Extras	2.7.2-1
Hive	Core	s3-dist-cp	Extras	2.36.0-1.amzn2023
Phoenix	Core	s3-dist-cp	Extras	2.36.0-1.amzn2023
Spark	Core	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Core	hadoop-mapreduce	Extras	3.4.0. amzn.2-1.amzn2023
Tez	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023
Tez	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Core	hadoop-client	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-hdfs	Extras	3.4.0.amzn.2-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-hdfs-datanode	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-kms	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-mapreduce	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.2-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	mariadb-connector-java	Extras	2.7.2-1
Trino	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Core	ZooKeeper	Extras	3.9.2. amzn.0-1.amzn2023
Trino	Core	zookeeper-server	Extras	3.9.2. amzn.0-1.amzn2023
Trino	Core	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Core	hadoop	Extras	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	Extras	3.1.3. amzn.15-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
ZooKeeper	Core	emrfs	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-client	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-hdfs	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-datanode	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-journalnode	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-namenode	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-httpfs	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-kms	Extras	3.9.2. amzn.0-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce	Extras	3.4.0.amzn.2-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce-historyserver	Extras	3.4.0.amzn.2-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
ZooKeeper	Core	hadoop-yarn	Extras	3.4.0.amzn.2-1.amzn2023
ZooKeeper	Core	hadoop-yarn-nodemanager	Extras	3.4.0.amzn.2-1.amzn2023
ZooKeeper	Core	hadoop-yarn-proxyserver	Extras	3.4.0.amzn.2-1.amzn2023
ZooKeeper	Core	hadoop-yarn-resourcemanager	Extras	3.4.0.amzn.2-1.amzn2023
ZooKeeper	Core	hadoop-yarn-timelineserver	Extras	3.4.0.amzn.2-1.amzn2023
ZooKeeper	Core	ranger-kms	Extras	2.0.0-1.amzn2023

notas de la versión 7.6.0

Las siguientes notas de la versión incluyen información sobre la versión 7.6.0 de Amazon EMR.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.6.0 application upgrades include HBase 2.6.1, Flink 1.20.0, Spark 3.5.3, and Trino 457.

Problemas conocidos

- Cuando el escalado gestionado está activado, el YARN ResourceManager (RM) sufre un grave problema de bloqueo que hace que deje de responder al pasar del estado de desmantelamiento al de desmantelamiento con operaciones simultáneas. Esto ocurre especialmente en entornos de clústeres grandes. Cuando esto ocurre, tanto las transiciones de estado de nodo como las transiciones de estado de aplicación no se procesan, lo que paraliza de manera efectiva la

funcionalidad principal del. ResourceManager Como solución alternativa, reinicie la JVM de RM mediante el siguiente comando en el nodo principal:

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

- Cuando se inicia un clúster de EMR con programadores YARN distintos del `CapacityScheduler` o `FairScheduler` (por ejemplo, el programador FIFO), la `YARNMemoryAvailablePercentage` métrica no se transfiere a. CloudWatch Este problema afecta a la reducción de escala en los clústeres habilitados para el escalado administrado. También puede afectar al escalado gestionado personalizado si depende de la métrica. `YARNMemoryAvailablePercentage` Como solución alternativa, utilice la tecla `CapacityScheduler` o`FairScheduler`.

Cambios, mejoras y problemas resueltos

- CloudWatch el agente es compatible con los clústeres de roles en tiempo de ejecución para EMR 7.6 y versiones posteriores. Para obtener más información, consulta el CloudWatch agente de Amazon.
- A partir de la versión 7.6.0 de Amazon Elastic Map Reduce (EMR), el conector del sistema de archivos S3A ahora admite el cifrado del lado del cliente de Amazon S3.
- A partir de EMR 7.6.0, TLS es compatible con los puntos finales. HMaster RegionServer Para obtener más información, consulte la siguiente tabla sobre el soporte de TLS con:. HBase HBase

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyu Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,'	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de septiembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

Versiones Java predeterminadas de 7.6.0

La versión 7.6 y posteriores de Amazon EMR incluyen Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla se muestran las versiones predeterminadas de Java para las aplicaciones de Amazon EMR 7.6.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
ZooKeeper	17, 11, 8

Versiones de los componentes 7.6.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1.31.0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.2.1-amzn-1	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.15.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.19,0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-notebook-env	1.18.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.36.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.15.0	Conector S3Select de EMR
emr-wal-cli	1.3.1	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.67,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.20.0-amzn-0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.20.0-amzn-0	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.4.0-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.4.0-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.4.0-amzn-2	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	3.4.0-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.4.0-amzn-2	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.4.0-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.4.0-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.4.0-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.4.0-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.4.0-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	2.6.1-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.6.1-amzn-0	Servicio para prestar servicio a una o más regiones. HBase
hbase-client	2.6.1-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.6.1-amzn-0	Servicio que proporciona un punto final HTTP para. RESTful HBase
hbase-thrift-server	2.6.1-amzn-0	Servicio que proporciona un punto final de Thrift para. HBase
hbase-operator-tools	2.6.1-amzn-0	Herramienta de reparación de clústeres de Apache. HBase
hcatalog-client	3.1.3-amzn-15	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-15	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-15	Punto final HTTP que proporciona una interfaz REST para. HCatalog

Componente	Versión	Descripción
hive-client	3.1.3-amzn-15	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-15	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-15	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-15	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.15.0-amzn-3	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.15.0-amzn-3	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.15.0-amzn-3	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.15.0-amzn-3	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.6.1-amzn-2	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	12.3.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.2.0	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.287-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.287-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.287-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	457-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	457-amzn-0	Service para ejecutar partes de una consulta.
trino-client	457-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.3-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.5.3-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.3-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.3-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	24.10.1-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-13	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
tez-on-worker	0.10.2-amzn-13	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.11.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.2-amzn-0	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.2-amzn-0	ZooKeeper cliente de línea de comandos.

clasificaciones de configuración de 7.6.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

clasificaciones emr-7.6.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.

Clasificaciones	Descripción	Acciones de reconfiguración
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j2	Cambie los valores en el archivo log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml del Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.

Clasificaciones	Descripción	Acciones de reconfiguración
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

7.6.0 registro de cambios

Registro de cambios de la versión 7.6.0 y de las notas de la versión

Date	Evento	Descripción
13-01-2020	Publicación de documentos	Las notas de la versión 7.6.0 de Amazon EMR se publicaro n por primera vez
10-01-10	Implementación completa	Amazon EMR 7.6.0 se ha implementado completam ente en todas las regiones compatibles
26 de diciembre de 2025	Versión inicial	Amazon EMR 7.6.0 se implementó por primera vez en las regiones comerciales iniciales

Amazon EMR versión 7.5.0

Ciclo de vida compatible con 7.5.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.5.0 de Amazon EMR.

Fase de soporte	Date
Fecha de lanzamiento inicial	21 de noviembre de 2024
Soporte estándar hasta	20 de noviembre de 2026

Fase de soporte	Date
Fin del soporte	21 de noviembre de 2026
Fin de vida útil	21 de noviembre de 2027

Para obtener una lista de vulnerabilidades y exposiciones críticas y muy comunes (CVEs) que no afectan a los clústeres de EMR con el software y las configuraciones recomendados, consulte 7.5.0 Vulnerabilidades y exposiciones comunes conocidas para los motores principales.

Componentes de la versión 7.5.0

Amazon EMR 7.5.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

Hue, Oozie, Pig, Presto, Zeppelin JupyterHub TensorFlow

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de la aplicación 7.5.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
AWS SDK for Java	2.28.8, 1.12.772	2.28.8, 1.12.772	2.25,70, 1.12.747	2.23.18, 1.12.705
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.17
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0
Flink	1.19.1	1.19.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.10	2.5.5	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-2	0.15.0-amzn-1	0.15.0-amzn-0	0.14.1-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0	1.5.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.2.0	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,287	0,285	0,285
Spark	3.5.2	3.5.2	3.5.1	3.5.1
Sqoop	-	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	446	446	442	436
Zeppelin	0.11.1	0.11.1	0.11.1	0.10.1
ZooKeeper	3.9.2	3.9.2	3.9.1	3.9.1

7.5.0 vulnerabilidades y exposiciones comunes

La siguiente tabla muestra todo lo CVEs que no afecta a los clústeres de EMR que se ejecutan en las configuraciones recomendadas de Amazon EMR 7.5.0. Amazon EMR depende del código abierto original para la disponibilidad de las correcciones y proporcionará la última versión estable como parte de los componentes del motor principal en un plazo de 90 días a partir de su verificación por parte de Amazon EMR.

Esta tabla se actualizó el 21 de noviembre de 2024.

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2021-37533	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37533
CVE-2018-14721	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2020-26939	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2024-47554	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47554
CVE-2021-38153	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-38153
CVE-2020-36188	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-10673	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-8908	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8908
CVE-2022-3509	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-41966	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-41966

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-3635	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2017-17485	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2023-25194	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
GHSA-58QW-P7QM-5RVh	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/GHSA-58qw-p7qm-5rvh
CVE-2023-51074	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-51074
CVE-2023-20883	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-20883
CVE-2022-42004	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2024-47561	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47561
CVE-2023-2976	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2023-33202	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-33202

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-14892	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2023-33201	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-10237
CVE-2015-5237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2024-36114	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-36114
CVE-2024-7254	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2024-23944	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-23944
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2023-52428	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2024-29857

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2024-30171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-30171
CVE-2022-41854	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-41854
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2018-7489	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-15250	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-9548	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2024-29025	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-29025
CVE-2024-36124	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-36124

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2013-4002	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2020-36186	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2022-45688	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202245688
CVE-2020-17521	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-17521
CVE-2024-8184	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2023-40167	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2022-25647	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2018-14719	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2024-29131	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-29131
CVE-2020-9546	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-1000180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2021-22570	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2022-46751	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202246751
CVE-2021-36373	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36373
CVE-2023-26048	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-30172
CVE-2022-36364	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202236364
CVE-2022-46337	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2019-16869	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-202238751	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202238751
CVE-2022-1415	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-1415
CVE-2021-22569	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-22569

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2024-35255	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-35255
CVE-2021-28170	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-28170
CVE-2022-40149	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2023-50868	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-2023-34454	high	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-202238750	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238750
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2024-25638	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25638
CVE-2024-29133	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-29133
CVE-2022-37865	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202237865
CVE-2018-19361	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-19362

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-10650	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2024-23454	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2020-24750	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2019-14439	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2023-34455	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2024-26308	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2023-0833	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

ID de CVE	Gravedad	URL de detalles del CVE
CVE-202238752	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238752
CVE-2020-1945	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2020-13936	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2020-10968	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-28052	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-43859
CVE-2021-27568	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2023-5072	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2024-21634	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-21634

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36180	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-29425	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-29425
CVE-2020-8840	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2019-16943	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2016-5018	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-34447
CVE-2024-39676	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-39676
CVE-2021-21290	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2022-39135	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-202239135
CVE-2019-16335	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2020-14062	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-15522
CVE-2019-14540	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2017-15095	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-12814	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12814

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-10672	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-24616	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-13949	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2018-12022	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2020-36189	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-12384
CVE-2023-26049	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2022-1471	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2022-40152	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240152

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-20863	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-11798
CVE-2019-20330	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2020-11619	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-202240151	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240151
CVE-2019-17267	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2018-10054	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-10054

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-14720	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-202224329	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-202224329
CVE-2020-36183	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2022-37866	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202237866
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2018-1313
CVE-2018-12023	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2024-38808	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2012-0881	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2012-0881

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-13956	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13956
CVE-202238749	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202238749
CVE-2017-7525	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-19360	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2021-37136	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2022-42889	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202242889
CVE-2021-36374	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13955
CVE-2022-25857	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202225857

Detalles de los componentes de la versión 7.5.0

Consulte la siguiente tabla para obtener más información sobre los paquetes Extras de Amazon EMR 7.5.0.

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Flink	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hudi	Extras	0.15.0.amzn.0-1.amzn2023
Flink	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Core	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
HBase	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
HCatalog	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
HCatalog	Core	mariadb-connector-java	Extras	
Hive	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Core	hive-hbase	Extras	3.1.3.amzn.12-1.amzn2023
Hive	Core	mariadb-connector-java	Extras	2.7.2-1
Hive	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
Phoenix	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
Spark	Core	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Tez	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	mariadb-connector-java	Extras	2.7.2-1
Trino	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	Extras	1.2.4-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
ZooKeeper	Core	emrfs	Extras	2.64.0-1.amzn2023
ZooKeeper	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
ZooKeeper	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	ranger-kms	Extras	2.0.0-1.amzn2023

notas de la versión 7.5.0

Las siguientes notas de la versión incluyen información sobre la versión 7.5.0 de Amazon EMR.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.5.0 application upgrades include Delta 3.2.1-amzn-0, HBase 2.5.10, Hudi 0.15.0-amzn-2, and Iceberg 1.6.1-amzn-1.
- Los registros de escritura anticipada (WAL) de EMR admiten la replicación entre clústeres. Para obtener más información, consulte Replicación entre clústeres EMR WAL.

Problemas conocidos

- Se ha eliminado Apache Sqoop de la versión 7.5 y versiones posteriores de Amazon EMR.
- Cuando se habilita el escalado administrado, el YARN ResourceManager (RM) sufre un grave problema de bloqueo que hace que deje de responder al pasar del estado de desmantelamiento al de desmantelamiento con operaciones simultáneas. Esto ocurre especialmente en entornos de clústeres grandes. Cuando esto ocurre, tanto las transiciones de estado de nodo como las transiciones de estado de aplicación no se procesan, lo que paraliza de manera efectiva la funcionalidad principal del. ResourceManager Como solución alternativa, reinicie la JVM de RM mediante el siguiente comando en el nodo principal:

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

- Cuando se inicia un clúster de EMR con programadores YARN distintos del `CapacityScheduler` o `FairScheduler` (por ejemplo, el programador FIFO), la `YARNMemoryAvailablePercentage` métrica no se transfiere a. CloudWatch Este problema afecta a la reducción de escala en los clústeres habilitados para el escalado administrado. También puede afectar al escalado gestionado personalizado si depende de la métrica. `YARNMemoryAvailablePercentage` Como solución alternativa, utilice la tecla `CapacityScheduler` o`FairScheduler`.

Cambios, mejoras y problemas resueltos

- Amazon EMR utiliza diferentes compilaciones de la TensorFlow biblioteca, en función de los tipos de instancias que elija para el clúster. Amazon EMR también admite clústeres con tipos TensorFlow de instancias aarch64 (instancias Graviton) de EMR-7.5.0 y versiones posteriores.

  La siguiente tabla muestra las compilaciones por tipo de instancia para EMR-7.5.0:

EC2 tipos de instancias	TensorFlow construir		
M5 y C5	Tensorflow 2.16.1 con optimización MKL de Intel		
P2, P4D, P5, G4DN, G5, G6 y GR6	Tensorflow 2.16.1 con CUDA 12.3, cuDNN 8.9.7.29		
P3, P3DN, G3 y G3S	Tensorflow 2.16.1 con CUDA 12.3, cuDNN 8.9.7.29 y NCCL 2.20.3-1 Nvidia NCCL solo está disponible en instancias P3. Contrato de licencia de usuario final (EULA): al utilizar componentes de Nvidia en Amazon EMR, aceptas los términos y condiciones descritos en el CLUF del producto.		
Todos los demás, incluido Graviton (EMR-7.5.0+)	Tensorflow 2.16.1		

- Un clúster Amazon EMR habilitado para AWS Lake Formation con FGAC (Fine-Grained Access Control) ahora puede admitir hasta 256 pasos simultáneos con la siguiente configuración de hdfs-site.xml:

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

-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyu Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,1	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de septiembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

Versiones Java predeterminadas de 7.5.0

La versión 7.5 y posteriores de Amazon EMR incluyen Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla se muestran las versiones predeterminadas de Java para las aplicaciones de Amazon EMR 7.5.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	8
ZooKeeper	17, 11, 8

Versiones de componentes 7.5.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1.31.0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.2.1-amzn-0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.14.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.18.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-notebook-env	1.16.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.35.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.14.0	Conector S3Select de EMR
emr-wal-cli	1.3.1	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.66,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.19.1-amzn-1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.19.1-amzn-1	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.4.0-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.4.0-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.4.0-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	3.4.0-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.4.0-amzn-1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.4.0-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.4.0-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.4.0-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.4.0-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.4.0-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	2.5.10-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.5.10-amzn-0	Servicio para prestar servicio a una o más regiones. HBase
hbase-client	2.5.10-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.5.10-amzn-0	Servicio que proporciona un punto final HTTP para. RESTful HBase
hbase-thrift-server	2.5.10-amzn-0	Servicio que proporciona un punto final de Thrift para. HBase
hbase-operator-tools	2.5.10-amzn-0	Herramienta de reparación de clústeres de Apache. HBase
hcatalog-client	3.1.3-amzn-14	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-14	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-14	Punto final HTTP que proporciona una interfaz REST para. HCatalog

Componente	Versión	Descripción
hive-client	3.1.3-amzn-14	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-14	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-14	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-14	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.15.0-amzn-2	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.15.0-amzn-2	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.15.0-amzn-2	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.15.0-amzn-2	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.6.1-amzn-1	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	12.3.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.2.0	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.287-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.287-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.287-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	446-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	446-amzn-0	Service para ejecutar partes de una consulta.
trino-client	446-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.2-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	3.5.2-amzn-1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.2-amzn-1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.2-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	24.08.1-amzn-1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con. GPUs
tensorflow	2.16.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-12	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
tez-on-worker	0.10.2-amzn-12	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.11.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.2-amzn-0	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.2-amzn-0	ZooKeeper cliente de línea de comandos.

clasificaciones de configuración de 7.5.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

clasificaciones emr-7.5.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.

Clasificaciones	Descripción	Acciones de reconfiguración
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j2	Cambie los valores en el archivo log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml del Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.

Clasificaciones	Descripción	Acciones de reconfiguración
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

7.5.0 registro de cambios

Registro de cambios de la versión 7.5.0 y de las notas de la versión

Date	Evento	Descripción
21-11-2020	Publicación de documentos	Las notas de la versión de Amazon EMR 7.5.0 se publicaron por primera vez
21-11-2021	Implementación completa	Amazon EMR 7.5.0 se ha implementado completamente en todas las regiones compatibles
13 de noviembre de 2022	Versión inicial	Amazon EMR 7.5.0 se implementó por primera vez en las regiones comerciales iniciales

Amazon EMR versión 7.4.0

Ciclo de vida compatible con 7.4.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.4.0 de Amazon EMR.

Fase de soporte	Date
Fecha de lanzamiento inicial	13 de noviembre de 2024
Soporte estándar hasta	12 de noviembre de 2026

Fase de soporte	Date
Fin del soporte	13 de noviembre de 2026
Fin de vida útil	13 de noviembre de 2027

Para obtener una lista de vulnerabilidades y exposiciones críticas y muy comunes (CVEs) que no afectan a los clústeres de EMR con el software y las configuraciones recomendados, consulte 7.4.0 Vulnerabilidades y exposiciones comunes conocidas para los motores principales.

Componentes de la versión 7.4.0

Amazon EMR 7.4.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

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

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de la aplicación 7.4.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,, Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, Oozie, Phoenix, Pig, Presto, Spark Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
AWS SDK for Java	2.28.8, 1.12.772	2.25,70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300032,2
Delta	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0	3.0.0
Flink	1.19.1	1.18.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.5	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-1	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-0	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,287	0,285	0,285	0,284
Spark	3.5.2	3.5.1	3.5.1	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	446	442	436	435
Zeppelin	0.11.1	0.11.1	0.10.1	0.10.1
ZooKeeper	3.9.2	3.9.1	3.9.1	3.9.1

7.4.0 vulnerabilidades y exposiciones comunes

La siguiente tabla muestra todo lo CVEs que no afecta a los clústeres de EMR que se ejecutan en las configuraciones recomendadas de Amazon EMR 7.4.0. Amazon EMR depende del código abierto original para la disponibilidad de las correcciones y proporcionará la última versión estable como parte de los componentes del motor principal en un plazo de 90 días a partir de su verificación por parte de Amazon EMR.

Esta tabla se actualizó el 15 de noviembre de 2024.

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-37533	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37533
CVE-2018-14721	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2020-26939	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-26939
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2024-47554	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47554
CVE-2021-38153	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-38153
CVE-2022-41966	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-41966
CVE-2023-3635	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2017-17485	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25710
CVE-2023-25194	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-25194

ID de CVE	Gravedad	URL de detalles del CVE
GHSA-58QW-P7QM-5RVh	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/GHSA-58qw-p7qm-5rvh
CVE-2023-51074	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-51074
CVE-2023-20883	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-20883
CVE-2022-42004	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2024-47561	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-47561
CVE-2023-2976	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2023-33202	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-33202
CVE-2019-14892	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2023-33201	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-33201
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-10237
CVE-2015-5237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-5237

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2024-36114	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-36114
CVE-2024-7254	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2024-23944	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-23944
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2023-52428	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2024-29857
CVE-2024-30171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-30171
CVE-2022-41854	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-41854
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2018-7489	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-15250	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-15250
CVE-2020-9548	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-51504
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2024-29025	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-29025
CVE-2024-36124	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-36124
CVE-2013-4002	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2020-36186	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2022-45688	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202245688
CVE-2020-17521	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-17521

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2024-8184	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2023-40167	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-6763
CVE-2022-25647	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2018-14719	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2024-29131	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-29131
CVE-2020-9546	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2018-1000180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2021-22570	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2021-22570
CVE-2022-46751	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202246751
CVE-2021-36373	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2021-36373

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-26048	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-30172
CVE-2022-36364	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202236364
CVE-2022-46337	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2019-16869	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-202238751	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202238751
CVE-2022-1415	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-1415
CVE-2021-22569	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-22569
CVE-2024-35255	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-35255
CVE-2021-28170	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-28170
CVE-2022-40149	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2023-50868	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-50868
CVE-202238750	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238750

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-42503
CVE-2024-25638	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-25638
CVE-2024-29133	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-29133
CVE-2022-37865	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202237865
CVE-2018-19361	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2020-10650	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	moderar	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2024-23454	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2020-24750	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-11112

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-14439	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2023-34455	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2024-26308	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2023-0833	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2012-5783	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2012-5783
CVE-202238752	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-202238752
CVE-2020-1945	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-1945
CVE-2020-13936	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2020-10968	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-28052	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-43859
CVE-2021-27568	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2023-5072	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202239135
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2017-7525	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2024-21634	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2012-6153	BAJA	https://nvd.nist. gov/vuln/detail/CVE-2012-6153
CVE-2020-36180	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-29425	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-29425
CVE-2020-8840	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2016-5018	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-34447
CVE-2024-39676	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-39676

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-21290	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2020-14062	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2020-15522
CVE-2019-14540	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2017-15095	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-12814	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12814

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-10672	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-24616	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-13949	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2018-12022	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2020-36189	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-12384
CVE-2023-26049	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2023-26049
CVE-2022-1471	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2022-40152	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240152

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-20863	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-11798
CVE-2019-20330	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2020-11619	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-202240151	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240151
CVE-2019-17267	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2018-10054	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-10054

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-14720	MEDIO	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-202224329	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-202224329
CVE-2020-36183	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2022-37866	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202237866
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2018-1313
CVE-2018-12023	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2024-38808	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
GHSA-CHFM-68VV-PVW5	ALTO	https://github.com/advisories/GHSA-CHFM-68VV-PVW5

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2012-0881	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-13956	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13956
CVE-202238749	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202238749
CVE-2021-37136	SIN CLASIFICAR	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2022-42889	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202242889
CVE-2021-36374	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13955
CVE-2022-25857	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202225857

Detalles de los componentes de la versión 7.4.0

Consulte la siguiente tabla para obtener más información sobre los paquetes Extras de Amazon EMR 7.4.0.

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Flink	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hudi	Extras	0.15.0.amzn.0-1.amzn2023
Flink	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Core	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
HBase	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
HCatalog	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
HCatalog	Core	mariadb-connector-java	Extras	
Hive	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Core	hive-hbase	Extras	3.1.3.amzn.12-1.amzn2023
Hive	Core	mariadb-connector-java	Extras	2.7.2-1
Hive	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
Phoenix	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
Spark	Core	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Tez	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	mariadb-connector-java	Extras	2.7.2-1
Trino	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	Extras	1.2.4-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
ZooKeeper	Core	emrfs	Extras	2.64.0-1.amzn2023
ZooKeeper	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
ZooKeeper	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	ranger-kms	Extras	2.0.0-1.amzn2023

notas de la versión 7.4.0

Las siguientes notas de la versión incluyen información sobre la versión 7.4.0 de Amazon EMR.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.4.0 application upgrades include AWS SDK for Java 2.28.8, Delta 3.2.0-amzn-1, Flink 1.19.1, Hadoop 3.4.0, HBase 2.5.5, Hudi 0.15.0-amzn-1, Iceberg 1.6.1-amzn-0, Phoenix 5.2.0, Presto 0.287, Spark 3.5.2, Trino 446, and Zookeeper 3.9.2.
- Esta versión agrega 7 puntos de enlace de código abierto que admiten el cifrado en tránsito a través de la red. Esta mejora incluye marcos de código abierto como Hue, Livy, Zookeeper y Ranger JupyterEnterpriseGateway KMS. Si ha activado el cifrado en tránsito y desea conectarse a estos puntos de conexión, es posible que tenga que modificar el código para que sea compatible con la comunicación TLS. Puede deshabilitar el TLS para cualquiera de estos puntos de conexión según sea necesario. Para obtener una lista completa de los puntos de conexión compatibles, consulte Descripción del cifrado en tránsito.

Problemas conocidos

- La `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` propiedad de Yarn ResourceManager, que anteriormente era compatible con los controladores de reproducción aleatoria de Hive desde la versión emr-6.11.0, ahora también permite retirar sin problemas los archivos servidos por el servicio de reproducción aleatoria externo de Spark cuando se ejecuta Spark en Yarn en emr-7.4.0.

  Tenga en cuenta que el comportamiento predeterminado del servicio aleatorio externo de Spark, cuando se ejecuta Spark en Yarn, es que Yarn elimine los archivos de reproducción aleatoria de las aplicaciones al finalizar la aplicación. NodeManager Esto podría afectar a la velocidad de desmantelamiento de los nodos y a la utilización del cómputo. En el caso de aplicaciones de ejecución prolongada, considere la posibilidad de configurar esta opción `spark.shuffle.service.removeShuffle true` para eliminar los archivos aleatorios que ya no estén en uso, lo que permite desmantelar más rápidamente los nodos sin datos de reproducción aleatoria activos.
- Disponibilidad de Amazon S3 Select: Amazon S3 Select no admite consultas entre regiones para Presto y Trino. Para obtener más información sobre las funciones específicas disponibles para las versiones de Presto y Trino, consulte Presto y Trino.
- La versión 7.4 de Amazon EMR no admite Apache Ranger.
- Apache Sqoop no es compatible con Apache Hadoop 3.4 debido a un problema conocido, el SQOOP-3485. La versión 7.4 de Amazon EMR incluye Apache Sqoop, pero no lo admite. Apache Sqoop se eliminará de la versión 7.5 y versiones posteriores de Amazon EMR.

- Cuando el escalado gestionado está activado, el YARN ResourceManager (RM) sufre un grave problema de bloqueo que hace que deje de responder al pasar del estado de desmantelamiento al de desmantelamiento con operaciones simultáneas. Esto ocurre especialmente en entornos de clústeres grandes. Cuando esto ocurre, tanto las transiciones de estado de nodo como las transiciones de estado de aplicación no se procesan, lo que paraliza de manera efectiva la funcionalidad principal del. ResourceManager Como solución alternativa, reinicie la JVM de RM mediante el siguiente comando en el nodo principal:

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

- Cuando se inicia un clúster de EMR con programadores YARN distintos del `CapacityScheduler` o `FairScheduler` (por ejemplo, el programador FIFO), la `YARNMemoryAvailablePercentage` métrica no se transfiere a. CloudWatch Este problema afecta a la reducción de escala en los clústeres habilitados para el escalado administrado. También puede afectar al escalado gestionado personalizado si depende de la métrica. `YARNMemoryAvailablePercentage` Como solución alternativa, utilice la tecla `CapacityScheduler` o`FairScheduler`.

Cambios, mejoras y problemas resueltos

- Amazon EMR utiliza diferentes compilaciones de la TensorFlow biblioteca en función de los tipos de instancias que elija para el clúster. Amazon EMR también admite clústeres con tipos TensorFlow de instancias aarch64 (instancias Graviton) de EMR-7.5.0 y versiones posteriores.
- Optimización del escalado de solicitudes de S3 para cargas de trabajo de Iceberg: para los clientes que utilizan Iceberg incluido en las versiones EMR 34 v7.4, S3 ha actualizado el esquema de entropía predeterminado a un hash de base 2 de 20 caracteres (por ejemplo, 35 01010110100110110010). Este nuevo esquema de entropía de base 2 reduce el rango de caracteres y permite a S3 escalar automáticamente la capacidad de solicitudes con mayor rapidez para adaptarla a las exigencias de la carga de trabajo de Iceberg objetivo. También reduce la cantidad de tiempo que una carga de trabajo observa 503 respuestas de aceleración.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Virginia del Norte), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,'	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241121.(	6.1.115-126.197. a.m. de 2023	12 de diciembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de septiembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

Versiones Java predeterminadas 7.4.0

La versión 7.4 y superior de Amazon EMR incluye Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla se muestran las versiones predeterminadas de Java para las aplicaciones de Amazon EMR 7.4.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	8
ZooKeeper	8

Versiones de los componentes 7.4.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que

pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1.31.0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.2.0-amzn-1	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.17.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.17.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.34.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.13.0	Conector S3Select de EMR
emr-wal-cli	1.3.1	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.65,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.19.1-amzn-0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.19.1-amzn-0	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.4.0-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.4.0-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.4.0-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.4.0-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.4.0-amzn-0	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.4.0-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.4.0-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.4.0-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.4.0-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.4.0-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.4.0-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.5.5-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.5.5-amzn-0	Servicio para prestar servicio a una o más regiones. HBase
hbase-client	2.5.5-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.5.5-amzn-0	Servicio que proporciona un punto final HTTP para. RESTful HBase
hbase-thrift-server	2.5.5-amzn-0	Servicio que proporciona un punto final de Thrift para. HBase
hbase-operator-tools	2.5.5-amzn-0	Herramienta de reparación de clústeres de Apache. HBase
hcatalog-client	3.1.3-amzn-13	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-13	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-13	Punto final HTTP que proporciona una interfaz REST para. HCatalog
hive-client	3.1.3-amzn-13	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-13	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-13	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-13	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.15.0-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.15.0-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.15.0-amzn-1	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.15.0-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.6.1-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	12.3.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.2.0	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.287-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.287-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.287-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	446-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	446-amzn-0	Service para ejecutar partes de una consulta.
trino-client	446-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.2-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.5.2-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.2-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.2-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	24.08.1-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con. GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.16.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.10.2-amzn-11	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-11	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.11.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.2-amzn-0	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.2-amzn-0	ZooKeeper cliente de línea de comandos.

clasificaciones de configuración 7.4.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

clasificaciones emr-7.4.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j2	Cambie los valores en el archivo log4j2.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml del Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

7.4.0 registro de cambios

Registro de cambios de la versión 7.4.0 y de las notas de la versión

Date	Evento	Descripción
18-11-2022	Publicación de documentos	Las notas de la versión 7.4.0 de Amazon EMR se publicaro n por primera vez
13-11-2020	Implementación completa	Amazon EMR 7.4.0 se ha implementado completam ente en todas las regiones compatibles
31 de octubre de 2022	Inicio del despliegue	Amazon EMR 7.4.0 se implementó por primera vez en las regiones comerciales iniciales

Amazon EMR, versión 7.3.0

Ciclo de vida compatible con 7.3.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.3.0 de Amazon EMR.

Fase de soporte	Date
Fecha de lanzamiento inicial	16 de octubre de 2024
Soporte estándar hasta	15 de octubre de 2026
Fin del soporte	16 de octubre de 2026
Fin de vida útil	16 de octubre de 2027

Para obtener una lista de vulnerabilidades y exposiciones críticas y muy comunes (CVEs) que no afectan a los clústeres de EMR con el software y las configuraciones recomendados, consulte 7.3.0 Vulnerabilidades y exposiciones comunes conocidas para los motores principales.

Componentes de la versión 7.3.0

Amazon EMR 7.3.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

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

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de las aplicaciones de la versión 7.3.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent, Delta, Flink, HBase,, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Pig, PrestoSpark Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18,1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

7.3.0 common vulnerabilities and exposures

La siguiente tabla muestra todo lo CVEs que no afecta a los clústeres de EMR que se ejecutan en las configuraciones recomendadas de Amazon EMR 7.3.0. Amazon EMR depende del código abierto original para la disponibilidad de las correcciones y proporcionará la última versión estable como parte de los componentes del motor principal en un plazo de 90 días a partir de su verificación por parte de Amazon EMR.

Esta tabla se actualizó el 4 de octubre de 2024.

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-1000180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-1000180
CVE-2020-28052	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-28052
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-40152	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240152
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-50387	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-50387
CVE-2012-0881	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12023

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-11112	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-35490	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36189	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35517

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-36090	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-36364	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202236364
CVE-2022-37865	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202237865
CVE-2022-37866	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202237866
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202239135
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2022-46751	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202246751

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2016-5018	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2015-2156	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-18640
CVE-2017-18640	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-18640
CVE-2019-16869	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2021-37136	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-36478	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2022-45688	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245688
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-5072	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

Detalles de los componentes de la versión 7.3.0

Consulte la siguiente tabla para obtener más información sobre los paquetes Extras de Amazon EMR 7.3.0.

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Flink	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Flink	Core	hudi	Extras	0.15.0.amzn.0-1.amzn2023
Flink	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Core	hadoop	Extras	3.3.6.amzn.5-1.amzn2023
HBase	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
HCatalog	Core	mariadb-connector-java	Extras	
Hive	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
Hive	Core	hive-hbase	Extras	3.1.3.amzn.12-1.amzn2023
Hive	Core	mariadb-connector-java	Extras	2.7.2-1
Hive	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
Phoenix	Core	s3-dist-cp	Extras	2.33.0-1.amzn2023
Spark	Core	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Tez	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
Trino	Core	hive	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	mariadb-connector-java	Extras	2.7.2-1
Trino	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Core	hadoop	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	Extras	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-jdbc	Extras	3.1.3.amzn.12-1.amzn2023
ZooKeeper	Core	emrfs	Extras	2.64.0-1.amzn2023
ZooKeeper	Core	hadoop-client	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-httpfs	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-kms	Extras	3.3.6.amzn.5-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
ZooKeeper	Core	hadoop-mapreduce	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.5-1.amzn2023
ZooKeeper	Core	ranger-kms	Extras	2.0.0-1.amzn2023

Notas de la versión 7.3.0

Las siguientes notas de la versión incluyen información sobre la versión 7.3.0 de Amazon EMR. Los cambios son respecto a la versión 7.3.0.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.3.0 application upgrades include Delta 3.2.0-amzn-0, Hudi 0.15.0-amzn-0, TensorFlow 2.16.1, Trino 442, and Zeppelin 0.11.1.

- Esta versión agrega 22 puntos de conexión de código abierto que admiten el cifrado en tránsito a través de la red. Esta mejora incluye marcos de código abierto como Hadoop, Hive, Tez HBase, Phoenix y Flink. Si ha activado el cifrado en tránsito y desea conectarse a estos puntos de conexión, es posible que tenga que modificar el código para que sea compatible con la comunicación TLS. Puede deshabilitar el TLS para cualquiera de estos puntos de conexión según sea necesario. Para obtener una lista completa de los puntos de conexión compatibles, consulte Descripción del cifrado en tránsito.
- Con Amazon EMR 7.3.0, Apache Hive with Iceberg ahora es compatible con el catálogo de datos de AWS Glue como metaalmacén, lo que le brinda ventajas como un repositorio centralizado para almacenar metadatos, rastrear automáticamente las fuentes de datos y extraer metadatos, y la integración con otros servicios para lograr flujos de trabajo sin problemas.

Problemas conocidos

- Python 3.11 no es compatible con EMR Studio.
- Si utiliza el cifrado en tránsito y la autenticación de Kerberos, no podrá utilizar las aplicaciones de Spark que dependan del metaalmacén de Hive. Este problema se solucionó en HIVE-16340 y se resolverá por completo en SPARK-44114 cuando Spark de código abierto pueda actualizarse a Hive 3. Mientras tanto, puede tratar de resolver este problema temporalmente si establece `hive.metastore.use.SSL` en `false`. Para obtener más información, consulte Configuración de aplicaciones.
  - Si no usa la autenticación de Kerberos, debe configurar Spark para que use SSL. Para ello, establezca las siguientes configuraciones cuando aprovisione el clúster.

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

- Si utilizas dominios personalizados, es posible que algunos enlaces de la interfaz de usuario de Spark History Server de la consola no funcionen. AWS Cargue el siguiente script en un bucket de S3 y, a continuación, añada un paso para usar el script del intérprete de comandos.

```
#!/bin/bash
```

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

sudo systemctl restart nginx
```

Cambios, mejoras y problemas resueltos

- Para permitir la verificación del nombre de host TLS para YARN, esta versión cambia la configuración de `yarn.resourcemanager.hostname` de la dirección IP de la instancia principal al nombre DNS privado de la instancia principal.
- MXNet El proyecto ha llegado al final de su vida útil y ahora está archivado. the section called “emr-7.2.0”es la última versión que incluye MXNet.

- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Virginia del Norte), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Oeste (Calgary), Europa

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,'	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241121.(	6.1.115-126.197. a.m. de 2023	12 de diciembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241031,(	6.1.112-124.190. amzn2023	15 de noviembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240916	6.1.109-118.189.am zn2023	6 de septiembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

Versiones predeterminadas de Java para la versión 7.3.0

Las versiones 7.3 y posteriores de Amazon EMR vienen con Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 7.3.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	8
Pig	8
ZooKeeper	17, 11, 8

Versiones de los componentes de la versión 7.3.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que

pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1,31,0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.2.0-amzn-0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.12.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.16.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.16.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.33.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.12.0	Conector S3Select de EMR
emr-wal-cli	1.3.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.64,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.18.1-amzn-2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.18.1-amzn-2	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.3.6-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.6-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.3.6-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.6-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.6-amzn-5	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.6-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.6-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.6-amzn-5	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.6-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.3.6-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-7	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.17-amzn-7	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-7	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-7	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-7	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-7	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-12	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-12	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-12	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-12	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-12	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-12	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-12	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.15.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.15.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.15.0-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.15.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.5.2-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	12.3.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.285-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.285-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.285-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	442-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	442-amzn-0	Service para ejecutar partes de una consulta.
trino-client	442-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.1-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	3.5.1-amzn-1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.1-amzn-1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.1-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	24.06.1-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.16.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.10.2-amzn-10	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-10	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.11.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.1-amzn-1	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.1-amzn-1	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 7.3.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-7.3.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

Registro de cambios de la versión 7.3.0

Registro de cambios de la versión 7.3.0 y notas de la versión

Date	Evento	Descripción
17-10-24	Publicación de documentos	Primera publicación de las notas de la versión 7.3.0 de Amazon EMR
2024-10-16	Implementación completa	Amazon EMR 7.3.0 se ha implementado completam ente en todas las regiones compatibles
2024-10-14	Versión inicial	Implementación inicial de Amazon EMR 7.3.0 en regiones comerciales limitadas

Amazon EMR, versión 7.2.0

Ciclo de vida compatible con 7.2.0

En la siguiente tabla se describen las fechas de ciclo de vida admitidas para la versión 7.2.0 de Amazon EMR.

Fase de soporte	Date
Fecha de lanzamiento inicial	25 de julio de 2024
Soporte estándar hasta	24 de julio de 2026
Fin del soporte	25 de julio de 2026
Fin de vida útil	25 de julio de 2027

Para obtener una lista de vulnerabilidades y exposiciones críticas y muy comunes (CVEs) que no afectan a los clústeres de EMR con el software y las configuraciones recomendados, consulte 7.2.0 Vulnerabilidades y exposiciones comunes conocidas para los motores principales. Para ver una lista de las CVEs que se corrigieron en la versión 7.2.0, consulte la sección 7.2.0 que corrigió las vulnerabilidades y exposiciones más comunes.

Componentes de la versión 7.2.0

Amazon EMR 7.2.0 incluye las siguientes aplicaciones:

Motores principales:

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

Extras:

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

Para obtener más información sobre los paquetes incluidos, consulte Versiones de aplicaciones en las versiones 7.x de Amazon EMR. Para obtener más información sobre los motores principales y los extras, consulte Soporte estándar de Amazon EMR.

Versiones de las aplicaciones de la versión 7.2.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent, Delta, Flink, HBase,, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, PigPresto Spark, Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18,1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

Vulnerabilidades y exposiciones comunes de la versión 7.2.0

En esta página se enumeran todas las vulnerabilidades y exposiciones (CVEs) más comunes relacionadas con Amazon EMR 7.2.0.

Vulnerabilidades y exposiciones comunes conocidas de los motores principales con la versión 7.2.0

La siguiente tabla muestra todo lo CVEs que no afecta a los clústeres de EMR que se ejecutan en las configuraciones recomendadas de Amazon EMR 7.2.0. Amazon EMR depende del código abierto original para la disponibilidad de las correcciones y proporcionará la última versión estable como parte de los componentes del motor principal en un plazo de 90 días a partir de su verificación por parte de Amazon EMR.

Esta tabla se actualizó el 16 de septiembre de 2024.

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-40152	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240152
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2015-5237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485
CVE-2017-7525	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14718

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-10172	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-10172
CVE-2019-10202	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-10202
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35490

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2021-33813	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-33813
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-25194	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34610	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34610
CVE-2023-3635	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2012-0881	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2012-0881

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2013-4002	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19362

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-10650	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13936
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-35490	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36183	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-31684	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-31684
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-36364	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202236364
CVE-2022-37865	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202237865

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-37866	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202237866
CVE-2022-39135	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202239135
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2022-46751	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202246751
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42889	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202242889
CVE-2022-46751	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202246751

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-20883	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-20883
CVE-2023-25194	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-5072

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-8768	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-8768
CVE-2021-32797	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-32797
CVE-2021-32798	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-32798
CVE-2022-24758	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202224758
CVE-2023-24580	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-24580
CVE-2023-36053	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-36053
CVE-2023-41164	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-41164
CVE-2023-43665	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43665
CVE-2023-46695	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-46695
CVE-2024-24680	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-24680
CVE-2016-5018	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2012-0881	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2013-4002
CVE-2019-17566	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-17566
CVE-2019-17566	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-17566
CVE-2020-11988	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11988
CVE-2021-33813	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-33813
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-34169	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202234169
CVE-2022-40146	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240146
CVE-2022-41704	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-41704

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-42890	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42890
CVE-2022-44729	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202244729
CVE-2022-44729	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202244729
CVE-2022-44729	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202244729
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2023-4759	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-4759
CVE-2015-2156	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-18640
CVE-2017-18640	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-18640
CVE-2019-16869	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2021-37136	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37136

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-37137	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-36478

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2015-2080	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-2080
CVE-2015-4852	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-4852
CVE-2015-6420	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-6420
CVE-2015-7501	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2015-7501
CVE-2017-1000034	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-1000034
CVE-2017-12612	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-12612
CVE-2017-15288	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-15288
CVE-2017-5637	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-5637
CVE-2017-7656	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-7656
CVE-2017-7657	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7657
CVE-2017-7658	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-7658

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2017-7687	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-7687
CVE-2017-9735	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-9735
CVE-2017-9790	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2017-9790
CVE-2018-11793	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-11793
CVE-2018-11804	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-11804
CVE-2018-17190	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-17190
CVE-2018-8012	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-8012
CVE-2019-0204	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0204
CVE-2019-10099	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-10099
CVE-2020-17533	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-17533
CVE-2020-9480	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9480
CVE-2021-28165	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-28165
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-35517	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2021-36090	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-37866	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202237866
CVE-2022-46751	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202246751

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2014-0114	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2014-0114
CVE-2016-1000338	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-1000338
CVE-2016-1000340	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-1000340
CVE-2016-1000342	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-1000342
CVE-2016-1000343	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-1000343
CVE-2016-1000344	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-1000344
CVE-2016-1000352	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2016-1000352

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2017-1000487	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-1000487
CVE-2017-12629	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-12629
CVE-2018-10936	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-10936
CVE-2018-3258	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-3258
CVE-2018-3827	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-3827
CVE-2018-3831	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-3831
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2019-10086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-10086
CVE-2019-10086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-10086
CVE-2019-7611	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-7611
CVE-2020-13692	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13692
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2020-28491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-28491

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-7692	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-7692
CVE-2021-22573	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-22573	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-25646	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-25646
CVE-2021-26291	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-26291
CVE-2021-26291	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-26291
CVE-2021-26919	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-26919
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-1471	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-1471
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-25857	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225857
CVE-2022-31197	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-31197
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-33980	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202233980
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-4244	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-4244
CVE-2022-45688	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245688

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-1370	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1370
CVE-2023-1428	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1428
CVE-2023-25194	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-31418	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-31418
CVE-2023-32731	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-32731
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-39410	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-44981
CVE-2023-46120	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-46120
CVE-2023-5072	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2024-1597	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2021-35958	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-35958
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949
CVE-2021-34538	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-34538

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-45688	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245688
CVE-2022-46337	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202246337
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-5072	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2024-21634
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-36090
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2014-0114	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2014-0114
CVE-2015-2156	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2015-2156
CVE-2018-10054	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-10054
CVE-2019-0205	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-0205
CVE-2019-10086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-10086
CVE-2019-10086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-10086
CVE-2019-16869	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-16869
CVE-2019-17195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17195
CVE-2019-20444	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20445
CVE-2020-13949	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-13949

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-9492	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-9492
CVE-2021-37136	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37137
CVE-2021-37404	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2021-37404
CVE-2022-25168	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-5168
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-26612	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-26612
CVE-2022-3171	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	ALTO	https://nvd.nist. gov/vuln/detail/CVE-20223510
CVE-2022-40152	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240152
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-3635	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2023-38647
CVE-2023-43642	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-43642
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2024-1597
CVE-2017-15095	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2017-17485

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2018-11307	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-12086	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14379	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14439	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14540	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14892	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16335	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-16942	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-16943	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-16943
CVE-2019-17195	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17195
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17267	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-17531	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2019-20330	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10650	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-10672	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10673	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10968	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-10969	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11111	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-11620	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14060	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14061	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14062	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-14195	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24616	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750
CVE-2020-24750	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-24750

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-25649	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-25649	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-25649	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-25649
CVE-2020-35211	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35211
CVE-2020-35490	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35491	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-35728

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36180	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36181	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36183	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36183	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36184	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-36187	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36188	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36188	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2020-36518
CVE-2020-7692	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-7692
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2020-8840	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9546	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9547	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2020-9548	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-20190	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-20190
CVE-2021-22573	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-22573
CVE-2021-33036	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-33036
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2021-35515	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-35516
CVE-2021-37714	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-37714
CVE-2021-46877	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-46877
CVE-2022-25647	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202225647
CVE-2022-32532	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-2022-32532
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-40664	CRÍTICO	https://nvd.nist. gov/vuln/detail/CVE-202240664

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45688	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245688
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2023-2976	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-2976
CVE-2022-40149	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240149
CVE-2022-40150	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202240150
CVE-2022-45685	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245685
CVE-2022-45693	ALTO	https://nvd.nist. gov/vuln/detail/CVE-202245693
CVE-2023-1436	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6378	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2023-6481

Vulnerabilidades y exposiciones comunes corregidas de los motores principales con la versión 7.2.0

En la siguiente tabla se enumeran todas las correcciones CVEs en Amazon EMR 7.2.0.

ID de CVE	Gravedad	URL de detalles del CVE
CVE-2019-12402	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2019-12402
CVE-2021-41561	ALTO	https://nvd.nist. gov/vuln/detail/CVE-2021-41561

Detalles de los componentes de la versión 7.2.0

Consulte la siguiente tabla para obtener más información sobre los paquetes Extras de Amazon EMR 7.2.0.

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	Extras	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	Extras	2.32.0-1.amzn2023
AmazonCloudWatchAgent	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-hdfs	Extras	3.3.6.amzn.4-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Flink	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-httpfs	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-kms	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.4-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Flink	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.4-1.amzn2023
Flink	Core	hudi	Extras	0.14.1.amzn.1-1.amzn2023
Flink	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Flink	Core	hadoop	Extras	3.3.6.amzn.4-1.amzn2023
HBase	Core	s3-dist-cp	Extras	2.32.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
HCatalog	Core	mariadb-connector-java	Extras	2.7.2-1
Hive	Core	hadoop-httpfs	Extras	3.3.6.amzn.4-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Hive	Core	hive-hbase	Extras	3.1.3.amzn.11-1.amzn2023
Hive	Core	mariadb-connector-java	Extras	2.7.2-1

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Hive	Core	s3-dist-cp	Extras	2.32.0-1.amzn2023
Phoenix	Core	s3-dist-cp	Extras	2.32.0-1.amzn2023
Spark	Core	spark-datanucleus	Extras	3.5.1.amzn.0-1.amzn2023
Tez	Core	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
Tez	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023
Tez	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	bigtop-utils	Extras	1.2.0-1.amzn2023
Trino	Core	hadoop-client	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-hdfs	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.4-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	hadoop-kms	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hive	Extras	3.1.3.amzn.11-1.amzn2023
Trino	Core	hive-hcatalog-server	Extras	3.1.3.amzn.11-1.amzn2023
Trino	Core	mariadb-connector-java	Extras	2.7.2-1

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
Trino	Core	ranger-kms	Extras	2.0.0-1.amzn2023
Trino	Core	ZooKeeper	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	Extras	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	Extras	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	Extras	1.2.4-1.amzn2023
Trino	Core	hadoop	Extras	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-lzo	Extras	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	Extras	3.1.3.amzn.11-1.amzn2023
Trino	Core	hive-jdbc	Extras	3.1.3.amzn.11-1.amzn2023
ZooKeeper	Core	emrfs	Extras	2.63.0-1.amzn2023
ZooKeeper	Core	hadoop-client	Extras	3.3.6.amzn.4-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
ZooKeeper	Core	hadoop-hdfs	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-datanode	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-journalnode	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-hdfs-namenode	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-httpfs	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-kms	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-mapreduce-historyserver	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-yarn	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-yarn-nodemanager	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-yarn-proxyserver	Extras	3.3.6.amzn.4-1.amzn2023

Aplicación	Tipo de aplicación	Paquete	Tipo de paquete	Versión de paquete
ZooKeeper	Core	hadoop-yarn-resourcemanager	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	hadoop-yarn-timelineserver	Extras	3.3.6.amzn.4-1.amzn2023
ZooKeeper	Core	ranger-kms	Extras	2.0.0-1.amzn2023

Notas de la versión 7.2.0

Las siguientes notas de la versión incluyen información sobre la versión 7.2.0 de Amazon EMR. Los cambios son respecto a la versión 7.2.0.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.2.0 application upgrades include Iceberg 1.5.0-amzn-0 and Delta 3.1.0.
- Amazon EMR añade compatibilidad para que pueda utilizar otras aplicaciones HBase, como Flink y Hive, con la clase de almacenamiento Amazon S3 Express One Zone.
- Esta versión añade la capacidad de leer los objetos restaurados, de modo que puede leer los objetos de Glacier desde una ubicación S3 con el protocolo `S3A`. Esta funcionalidad funciona con Spark, Flink y Hive.
- Etiquetas de nodos con escalado gestionado: utilice etiquetas de nodos con sus instancias según el tipo de mercado o el tipo de nodo para mejorar el escalado automático desde Amazon EMR. Para obtener más información, consulte Uso del escalado automático en Amazon EMR.

Problemas conocidos

- Python 3.11 no es compatible con EMR Studio.

Cambios, mejoras y problemas resueltos

- Esta versión corrige un problema de bloqueo que se puede producir durante las operaciones internas de limpieza por pasos. Esta operación administra el ciclo de vida de los pasos a medida que se completan en el clúster de EMR. Este problema afecta a las operaciones críticas de Amazon EMR, como la operación escalonada y el escalado.
- Esta versión resuelve un problema por el que los clústeres personalizados con archivos de registro personalizados AMIs que tienen determinados archivos de registro preexistentes podían provocar un error en el daemon de administración de registros de Amazon EMR.
- Amazon EMR 7.2.0 actualiza el daemon de Amazon EMR responsable de las actividades de administración y supervisión de clústeres de la versión 1 a la versión 2 del SDK. AWS

- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Virginia del Norte), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,'	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241121.(	6.1.115-126.197. a.m. de 2023	12 de diciembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241031,(	6.1.112-124.190. amzn2023	15 de noviembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 241001.1	6.1.109-118.189.am zn2023	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240819,(	6.1.102-111.182.am zn2023	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240730,(	6.1.97-104.177.amzn2023	2 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240722,(	6.1.97-104.177.amzn2023	24 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240708.(	6.1.96-102.177.amzn2023	8 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

Versiones predeterminadas de Java para la versión 7.2.0

Las versiones 7.2 y posteriores de Amazon EMR vienen con Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 7.2.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	21, 17
Zeppelin	8
Pig	8
ZooKeeper	8

Versiones de los componentes de la versión 7.2.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que

pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1,31,0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.1.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.4.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.11.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.15.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.16.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.32.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.11.0	Conector S3Select de EMR
emr-wal-cli	1.2.2	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.63,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.18.1-amzn-1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.18.1-amzn-1	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.3.6-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.6-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.3.6-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.6-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.6-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.6-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.6-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.6-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.6-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.3.6-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-6	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.17-amzn-6	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-6	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-6	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-6	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-6	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-11	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-11	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-11	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-11	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-11	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-11	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-11	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.14.1-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.14.1-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.14.1-amzn-1	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.14.1-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.5.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-connectors	5.1.3	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.285-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.285-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.285-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	436-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	436-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
trino-client	436-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.1-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.5.1-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.1-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.1-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	24.02.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-9	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-9	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.1-amzn-1	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.1-amzn-1	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 7.2.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-7.2.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.	Restarts Zookeeper server.
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.	Restarts the CloudWatc hAgent service.

Registro de cambios de la versión 7.2.0

Registro de cambios de la versión 7.2.0 y notas de la versión

Date	Evento	Descripción
2024-07-25	Publicación de documentos	Primera publicación de las notas de la versión 7.2.0 de Amazon EMR
2024-07-16	Versión inicial	Implementación inicial de Amazon EMR 7.2.0 en regiones comerciales limitadas

Versión 7.1.0 de Amazon EMR

Versiones de las aplicaciones de la versión 7.1.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,,, Delta,, Flink, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue,, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNet, OoziePhoenixPigPresto Spark, Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18,1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,285	0,285	0,284	0,283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

Notas de la versión 7.1.0

Las siguientes notas de la versión incluyen información sobre la versión 7.1.0 de Amazon EMR. Los cambios son respecto a la versión 7.1.0.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.1.0 application upgrades include Livy 0.8.0, Trino 435, and ZooKeeper 3.9.1.
- Reemplazo de nodos en mal estado: con Amazon EMR 7.1.0 y versiones posteriores, el reemplazo de nodos en mal estado está habilitado de forma predeterminada, por lo que Amazon EMR reemplazará sin problemas los nodos en mal estado. Para evitar que sus flujos de trabajo actuales se vean afectados en las versiones 7.0.0 y anteriores de Amazon EMR, el reemplazo de nodos en mal estado está deshabilitado si ha activado la protección de terminación en su clúster.
- CloudWatch Agente: configure el CloudWatch agente para que utilice métricas de sistema adicionales, añada métricas de aplicación y cambie el destino de las métricas con la API de configuración de Amazon EMR.

Problemas conocidos

- La aplicación JupyterEnterpriseGateway (JEG) no es compatible con Python 3.11 en esta versión.

Cambios, mejoras y problemas resueltos

- Si bien Amazon EMR 7.1.0 es compatible con Python 3.9 de forma predeterminada, Livy 0.8.0 y Spark de Amazon EMR 7.1.0 admiten Python 3.11.
- Esta versión corrige el problema de tener que ejecutar cada línea de una en una cuando se usa PySpark con Python versión 3.11.
- Actualización de Zeppelin: Amazon EMR 7.1.0 incluye una actualización de Zeppelin al SDK para Java v2. AWS Esta actualización permite que un portátil Zeppelin S3 acepte un proveedor de materiales de cifrado personalizados. El AWS SDK for Java v2 elimina la EncryptionMaterialsProvider interfaz. Al actualizar a Amazon EMR 7.1.0, debe implementar la interfaz Keyring si desea utilizar un cifrado personalizado. Para ver un ejemplo de cómo implementar la interfaz Keyring, consulte KmsKeyring.java.
- Al actualizar a la versión 7.1.0 de Amazon EMR, cambie su proveedor de claves personalizado para el cifrado de discos locales a fin de generar claves mediante el algoritmo AES de. AES/GCM/NoPadding Si no actualiza el algoritmo, es posible que se produzca un error `Local disk`

`encryption failed on master instance (i-123456789) due to internal error.` en la creación del clúster. Para obtener más información sobre la creación de un proveedor de claves personalizadas, consulte Creación de un proveedor de claves personalizadas.

- Amazon EMR 7.1.0 mejora la resiliencia de un nodo en condiciones de poco espacio en disco al mejorar la lógica de truncamiento de registros para los archivos con identificadores de archivos abiertos.
- Esta versión mejora la lógica de codificación y decodificación para minimizar el riesgo de corrupción de datos y errores en los nodos cuando los daemons de Amazon EMR leen y escriben archivos al reiniciar un nodo.

- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Virginia del Norte), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250512.(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,'	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241121.(	6.1.115-126.197. a.m. de 2023	12 de diciembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241031,(	6.1.112-124.190. amzn2023	15 de noviembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 241001.1	6.1.109-118.189.am zn2023	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240819,(	6.1.102-111.182.am zn2023	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240730,(	6.1.97-104.177.amzn2023	2 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240722,(	6.1.97-104.177.amzn2023	24 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240708.(	6.1.96-102.177.amzn2023	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.3.2 240219,(	6.1.77-99.164.amzn2023	8 de mayo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

Versiones predeterminadas de Java para la versión 7.1.0

Las versiones 7.1 y posteriores de Amazon EMR vienen con Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 7.1.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
ZooKeeper	8

Versiones de los componentes de la versión 7.1.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que

pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1,31,0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.0.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.10.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.14.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.31.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.10.0	Conector S3Select de EMR
emr-wal-cli	1.2.2	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.62.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.18.1-amzn-0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.18.1-amzn-0	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.3.6-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.6-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.3.6-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.6-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.6-amzn-3	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.6-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.6-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.6-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.6-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.3.6-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-5	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.17-amzn-5	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-5	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-5	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-5	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-5	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-10	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-10	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-10	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-10	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-10	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-10	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-10	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.14.1-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.14.1-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.14.1-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.14.1-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.4.3-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.8.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-connectors	5.1.3	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.284-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.284-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.284-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	435-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	435-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
trino-client	435-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.0-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	3.5.0-amzn-1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.0-amzn-1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.0-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.10.0-amzn-1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-8	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-8	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.9.1-amzn-0	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.9.1-amzn-0	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 7.1.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-7.1.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
docker-conf	Cambie la configuración relacionada con el docker.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/Yarn
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-env	Cambiar los valores en el entorno de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive	Configuración de Amazon EMR mantenida para Apache Hive.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
hudi-defaults	Cambie los valores en el archivo hudi-defa ults.conf de Hudi.
iceberg-defaults	Cambie los valores del archivo iceberg-d efaults.conf de Iceberg.
delta-defaults	Cambie los valores del archivo delta-def aults.conf de Delta.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.

Clasificaciones	Descripción
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.

Clasificaciones	Descripción
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
trino-log	Cambiar los valores en el archivo log.properties de Trino.
trino-config	Cambiar los valores en el archivo config.properties de Trino.
trino-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Trino.
trino-env	Cambie los valores del archivo trino-env.sh de Trino.
trino-node	Cambiar los valores en el archivo node.properties de Trino.

Clasificaciones	Descripción
trino-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Trino.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Trino.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.
trino-connector-hive	Cambiar los valores en el archivo hive.prop erties de Trino.
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.p roperties de Trino.
trino-connector-hudi	Cambiar los valores en el archivo hudi.prop erties de Trino.
trino-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Trino.
trino-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Trino.
trino-connector-localfile	Cambiar los valores en el archivo localfile .properties de Trino.
trino-connector-memory	Cambiar los valores en el archivo memory.pr operties de Trino.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Trino.
trino-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Trino.

Clasificaciones	Descripción
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-logback	Cambie los valores en ZooKeeper el archivo logback.xml.

Clasificaciones	Descripción
emr-metrics	Cambie la configuración de la métrica EMR para este nodo.

Registro de cambios de la versión 7.1.0

Registro de cambios de la versión 7.1.0 y notas de la versión

Date	Evento	Descripción
5-10 de abril de 2022	Publicación de documentos	Primera publicación de las notas de la versión 7.1.0 de Amazon EMR
2024-05-08	Implementación completa	Amazon EMR 7.1.0 se ha implementado completam ente en todas las regiones compatibles
2024-04-23	Versión inicial	Implementación inicial de Amazon EMR 7.1.0 en regiones comerciales limitadas

Amazon EMR, versión 7.0.0

Versiones de las aplicaciones de la versión 7.0.0

Esta versión incluye las siguientes aplicaciones: AmazonCloudWatchAgent,,, Delta,, Flink, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue,, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNet, OoziePhoenixPigPresto Spark, Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS SDK for Java	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032,2	1,300032,2	1,300032,2	1,300031,1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18,1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0

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

Notas de la versión 7.0.0

Las siguientes notas de la versión incluyen información sobre la versión 7.0.0 de Amazon EMR. Los cambios son respecto a la versión 6.15.0.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.0.0 application upgrades include Python 3.9, Spark 3.5, Flink 1.18, and Delta 3.0. Esta versión también agrega soporte para la aplicación CloudWatch Agente de Amazon y elimina el soporte para Ganglia.
- Amazon Corretto 17: las versiones 7.0 y posteriores de Amazon EMR vienen con Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy. Para obtener más información acerca de las versiones del JDK compatibles con las aplicaciones de esta versión, consulte Versiones predeterminadas de Java para la versión 7.0.0.
- Amazon Linux 2023: con la versión 7.0, los clústeres de Amazon EMR ahora se ejecutan en la versión AL2 023 de forma predeterminada. Para obtener información sobre cómo afecta esto a la versión de la AMI predeterminada, consulte Consideraciones sobre la actualización de software en la Guía de administración de Amazon EMR. Tenga en cuenta también que AL2 023 eliminó Python 2.7, por lo que cualquier componente que requiera Python ahora debería escribirse con Python 3.
- S3 en Outposts con **`s3a`**: Amazon EMR ahora es compatible con los buckets de Amazon S3 en Outposts con el sistema de archivos `s3a`. Para obtener más información acerca de S3 en Outposts, consulte ¿Qué es S3 en Outposts? en la Guía del usuario de Amazon Simple Storage Service.

Problemas conocidos

- Tenga en cuenta que no puede ejecutar más de una línea a la vez si utiliza Python PySpark con la versión 3.10 o posterior. Debe ejecutar cada línea de una en una.

Cambios, mejoras y problemas resueltos

- Para mantener el estado de todas las instancias de un clúster, Amazon EMR se integra con Apache YARN, Apache HDFS y Kerberos. Con la versión 7.0, mejoramos estas integraciones para eliminar de forma fiable el estado de las instancias que se terminan debido al escalado o a otras operaciones. Esto es especialmente importante en el caso de los clústeres de larga duración que tienen habilitado el escalado administrado, ya que se escalan con frecuencia y acumulan miles de instancias terminadas a lo largo de su vida útil.
- Esta versión mejora la configuración de Kerberos para incluir únicamente la compatibilidad con cifrados basados en AES. Los clústeres de EMR que se ejecutan en las versiones 7.0.0

y posteriores de Amazon EMR ya no admiten el KDC de Kerberos con cifrados que no estén basados en AES. Los cifrados basados en AES ofrecen la máxima seguridad para sus clústeres.

- Como parte de la migración del AWS SDK 2.x, Amazon EMR 7.0 incluye una actualización del conector Spark Kinesis por motivos de compatibilidad. Esta actualización no está disponible en la versión comunitaria de Apache Spark. Si utiliza el conector de Spark Kinesis de una versión de Amazon EMR anterior a la 7.0, debe migrar los códigos de sus aplicaciones para que se ejecuten en el SDK 2.x antes de poder migrar sus cargas de trabajo a Amazon EMR 7.0. Para obtener más información, consulte Migración del conector de Spark Kinesis al SDK 2.x para Amazon EMR 7.0.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este) AWS GovCloud , (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (Oregón), EE.UU. Oeste (N. California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
23.7.202 0512.0	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250428,'	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Centro), Canadá Oeste

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250414.(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
23.7.202 0331.0	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218,:	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241121.(	6.1.115-126.197. a.m. de 2023	12 de diciembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 241031,(	6.1.112-124.190. amzn2023	15 de noviembre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 241001.1	6.1.109-118.189.am zn2023	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240819,(	6.1.102-111.182.am zn2023	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240730,(	6.1.97-104.177.amzn2023	2 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240722,(	6.1.97-104.177.amzn2023	24 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.5.2 240708.(	6.1.96-102.177.amzn2023	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.3.2 240304.(	6.1.79-99.164.amzn2023	12 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.3.2 240219,(	6.1.77-99.164.amzn2023	1 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.3.2 240205.(	6.1.75-99.163.amzn2023	19 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.3.2 240122,(	6.1.72-96.166.amzn2023	5 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.3.2 240108,(	6.1.72-96.166.amzn2023	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.3.2 231211.4	6.1.66-91.160.amzn2023	19 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

Versiones predeterminadas de Java para la versión 7.0.0

Las versiones 7.0 y posteriores de Amazon EMR vienen con Amazon Corretto 17 (basado en OpenJDK) de forma predeterminada para las aplicaciones compatibles con Corretto 17 (JDK 17), con la excepción de Apache Livy.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 7.0.0. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
ZooKeeper	8

Versiones de los componentes de la versión 7.0.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
adot-java-agent	1,31,0	Un agente Java que recopila métricas de los daemons de las aplicaciones.
delta	3.0.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
delta-standalone-connectors	0.6.0	Los conectores de Delta ofrecen diferentes tiempos de ejecución para integrar Delta Lake con motores como Flink, Hive y Presto.
emr-amazon-cloudwatch-agent	1.300031.1-amzn-0	Aplicación que recopila métricas internas a nivel del sistema y métricas de aplicaciones personalizadas de las instancias de Amazon EC2 .
emr-ddb	5.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.9.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.13.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.30.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.9.0	Conector S3Select de EMR
emr-wal-cli	1.2.1	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.61.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.18.0-amzn-0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.18.0-amzn-0	Administración de recursos en nodos EMR para Apache Flink. JobManager
hadoop-client	3.3.6-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.6-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	3.3.6-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.6-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.6-amzn-2	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.6-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.6-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.6-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.6-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	3.3.6-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-4	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.17-amzn-4	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-4	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-4	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-4	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-4	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-9	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-9	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	3.1.3-amzn-9	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-9	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-9	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-9	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-9	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.14.0-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.14.0-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.14.0-amzn-1	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.14.0-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
iceberg	1.4.2-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-connectors	5.1.3	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.283-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.283-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.283-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	426-amzn-1	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	426-amzn-1	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
trino-client	426-amzn-1	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.3.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.5.0-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.5.0-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.5.0-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.5.0-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.10.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-7	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-7	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.58	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 7.0.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-7.0.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
docker-conf	Cambie la configuración relacionada con el docker.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/Yarn
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-env	Cambiar los valores en el entorno de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive	Configuración de Amazon EMR mantenida para Apache Hive.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
hudi-defaults	Cambie los valores en el archivo hudi-defa ults.conf de Hudi.
iceberg-defaults	Cambie los valores del archivo iceberg-d efaults.conf de Iceberg.
delta-defaults	Cambie los valores del archivo delta-def aults.conf de Delta.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.

Clasificaciones	Descripción
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.

Clasificaciones	Descripción
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
trino-log	Cambiar los valores en el archivo log.properties de Trino.
trino-config	Cambiar los valores en el archivo config.properties de Trino.
trino-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Trino.
trino-env	Cambie los valores del archivo trino-env.sh de Trino.
trino-node	Cambiar los valores en el archivo node.properties de Trino.

Clasificaciones	Descripción
trino-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Trino.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Trino.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.
trino-connector-hive	Cambiar los valores en el archivo hive.prop erties de Trino.
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.p roperties de Trino.
trino-connector-hudi	Cambiar los valores en el archivo hudi.prop erties de Trino.
trino-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Trino.
trino-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Trino.
trino-connector-localfile	Cambiar los valores en el archivo localfile .properties de Trino.
trino-connector-memory	Cambiar los valores en el archivo memory.pr operties de Trino.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Trino.
trino-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Trino.

Clasificaciones	Descripción
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Registro de cambios de la versión 7.0.0

Registro de cambios de la versión 7.0.0 y notas de la versión

Date	Evento	Descripción
29-12-2020	Implementación completa	Amazon EMR 7.0.0 se ha implementado completam ente en todas las regiones compatibles
2023-12-28	Publicación de documentos	Primera publicación de las notas de la versión 7.0.0 de Amazon EMR
2023-12-19	Versión inicial	Implementación inicial de Amazon EMR 7.0.0 en regiones comerciales limitadas

Versiones de lanzamiento de Amazon EMR 6.x

Esta sección contiene las versiones de las aplicaciones, las notas de la versión, las versiones de los componentes y las clasificaciones de configuración de cada versión de lanzamiento de Amazon EMR 6.x.

Al lanzar un clúster, puede elegir entre varias versiones de Amazon EMR. Esto le permite probar y utilizar versiones de la aplicación que cumplan con sus requisitos de compatibilidad. El número de versión se especifica con la etiqueta de versión. Las etiquetas de versión tienen el formato `emr-x.x.x` Por ejemplo, `emr-7.9.0`.

Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Para obtener una tabla completa de las versiones de las aplicaciones incluidas en cada versión de Amazon EMR 6.x, consulte Versiones de la aplicación en las versiones 6.x de Amazon EMR.

Temas

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Amazon EMR, versión 6.15.0
- Amazon EMR, versión 6.14.0
- Amazon EMR, versión 6.13.0
- Amazon EMR, versión 6.12.0
- Amazon EMR, versión 6.11.1
- Amazon EMR, versión 6.11.0
- Amazon EMR, versión 6.10.1
- Amazon EMR, versión 6.10.0
- Amazon EMR, versión 6.9.1
- Amazon EMR, versión 6.9.0
- Amazon EMR, versión 6.8.1
- Amazon EMR, versión 6.8.0
- Amazon EMR, versión 6.7.0
- Amazon EMR, versión 6.6.0
- Amazon EMR, versión 6.5.0
- Amazon EMR, versión 6.4.0
- Amazon EMR, versión 6.3.1
- Amazon EMR, versión 6.3.0
- Amazon EMR, versión 6.2.1
- Amazon EMR, versión 6.2.0
- Amazon EMR, versión 6.1.1
- Amazon EMR, versión 6.1.0
- Amazon EMR, versión 6.0.1
- Amazon EMR, versión 6.0.0

Versiones de la aplicación en las versiones 6.x de Amazon EMR

Para ver una tabla completa con las versiones de las aplicaciones disponibles en cada versión de Amazon EMR 6.x, abra Versiones de la aplicación en las versiones 6.x de Amazon EMR en su navegador.

Amazon EMR, versión 6.15.0

Versiones de las aplicaciones de la versión 6.15.0

Esta versión incluye las siguientes aplicaciones: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Oozie, Phoenix, PigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
AWS SDK for Java	2.20.160-amzn-0, 1.12.569	1.12.543	1.12,513	1.12.490
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.17	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.4.0
Flink	1.17.1	1.17.1	1.17.0	1.17.0

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.14.0-amzn-0	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.4.0-amzn-0	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,283	0,281	0,281	0,281
Spark	3.4.1	3.4.1	3.4.1	3.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	426	422	414	414
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.15.0

Las siguientes notas de la versión incluyen información sobre la versión 6.15.0 de Amazon EMR. Los cambios son respecto a la versión 6.14.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.15.0.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.
- Lanzamientos más rápidos para clústeres de EMR que se ejecutan en EC2: ahora es hasta un 35% más rápido lanzar Amazon EMR en un clúster. EC2 Con esta mejora, la mayoría de los clientes pueden lanzar sus clústeres en 5 minutos o menos.
- CodeWhisperer para EMR Studio: ahora puede utilizar Amazon con CodeWhisperer Amazon EMR Studio para obtener recomendaciones en tiempo real a medida que escribe código. JupyterLab CodeWhisperer puede completar sus comentarios, terminar líneas de código individuales, hacer line-by-line recomendaciones y generar funciones con formato completo.
- Tiempos de reinicio de tareas más rápidos con Flink: con Amazon EMR 6.15.0 y versiones posteriores, Apache Flink dispone de varios mecanismos nuevos para mejorar el tiempo de reinicio de los trabajos durante las operaciones de recuperación o escalado de tareas. Esto optimiza la velocidad de los gráficos de recuperación y reinicio de la ejecución para mejorar la estabilidad del trabajo.
- Control de acceso detallado y a nivel de tabla para formatos de tabla abierta: con Amazon EMR 6.15.0 y versiones posteriores, cuando ejecuta trabajos de Spark en Amazon EMR EC2 en clústeres que acceden a los datos del catálogo de datos de Glue, AWS Lake Formation puede

utilizarlos para aplicar permisos a nivel de tabla, fila, columna y celda en AWS tablas basadas en Hudi, Iceberg o Delta Lake.

- Actualización de Hadoop: Amazon EMR 6.15.0 incluye una actualización de Apache Hadoop a la versión 3.3.6. Hadoop 3.3.6 era la versión más reciente en el momento de la implementación de Amazon EMR 6.15, lanzada por Apache en junio de 2023. Las versiones anteriores de Amazon EMR (6.9.0 a 6.14.x) utilizaban Hadoop 3.3.3.

  La actualización incluye cientos de mejoras y correcciones, además de características que incluyen parámetros de nodos de datos reconfigurables, la opción `DFSAdmin` para iniciar operaciones de reconfiguración masiva en todos los nodos de datos activos y una API vectorizada que permite a los lectores más exigentes especificar varios rangos de lectura. Hadoop 3.3.6 también añade compatibilidad con HDFS APIs y semántica para su registro de escritura anticipada (WAL), de modo que HBase pueda ejecutarse en otras implementaciones de sistemas de almacenamiento. Para obtener más información, consulte los registros de cambios de las versiones 3.3.4, 3.3.5 y 3.3.6 en la documentación de Apache Hadoop.
- Soporte para AWS SDK for Java, versión 2: las aplicaciones de Amazon EMR 6.15.0 pueden usar las versiones 1.12.569 o 2.20.160 AWS del SDK para Java si la aplicación es compatible con la versión 2. El AWS SDK para Java 2.x es una importante reescritura del código base de la versión 1.x. Se basa en Java 8 y agrega varias características solicitadas con frecuencia. Entre estas se incluyen la compatibilidad con operaciones de E/S sin bloqueo y la capacidad de conectar una implementación HTTP diferente en tiempo de ejecución. Para obtener más información, incluida una guía de migración del SDK para Java v1 a la v2, consulte la guía AWS SDK para Java, versión 2.

Cambios, mejoras y problemas resueltos

- Para mejorar sus clústeres de EMR de alta disponibilidad, esta versión permite la conectividad con los daemons de Amazon EMR en un host local que utiliza puntos de conexión IPv6.
- Esta versión permite que TLS 1.2 se comunique con todos los nodos principales del clúster de alta disponibilidad ZooKeeper aprovisionado.
- Esta versión mejora la administración de los archivos de registro de ZooKeeper transacciones que se mantienen en los nodos principales para minimizar los escenarios en los que los archivos de registro se exceden de los límites e interrumpen las operaciones del clúster.
- Esta versión hace que la comunicación entre nodos sea más resiliente para los clústeres de EMR de alta disponibilidad. Esta mejora reduce la posibilidad de que se produzcan errores en las acciones de arranque o en el inicio del clúster.

- En Amazon EMR 6.15.0, Tez presenta configuraciones que puede especificar para abrir de forma asíncrona las divisiones de entrada en una división agrupada de Tez. Esto se traduce en un rendimiento más rápido de las consultas de lectura cuando hay un gran número de divisiones de entrada en una sola división agrupada de Tez. Para obtener más información, consulte Apertura de división asíncrona de Tez.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich), Asia Pacífico (Malasia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	13 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

Versiones predeterminadas de Java para la versión 6.15.0

Las versiones 6.12.0 y posteriores de Amazon EMR son compatibles con todas las aplicaciones con Amazon Corretto 8 de forma predeterminada, excepto Trino. En el caso de Trino, Amazon EMR es compatible con Amazon Corretto 17 de forma predeterminada a partir de la versión 6.9.0 de Amazon EMR. Amazon EMR también admite algunas aplicaciones con Amazon Corretto 11 y 17. Esas aplicaciones se indican en la siguiente tabla. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Si bien Amazon EMR es compatible con Amazon Corretto 11 y 17 en Apache Spark, Apache Hadoop y Apache Hive, es posible que el rendimiento de algunas cargas de trabajo disminuya cuando utilice estas versiones de Corretto. Se recomienda que pruebe sus cargas de trabajo antes de cambiar los valores predeterminados.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 6.15.0:

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
ZooKeeper	8

Versiones de los componentes de la versión 6.15.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.4.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
delta-standalone-connectors	0.6.0	Los conectores de Delta ofrecen diferentes tiempos de ejecución para integrar Delta Lake con motores como Flink, Hive y Presto.
emr-ddb	5.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	3.8.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.12.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.29.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.8.0	Conector S3Select de EMR
emr-wal-cli	1.2.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.60.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.17.1-amzn-1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.17.1-amzn-1	Administración de recursos en nodos EMR para Apache Flink. JobManager

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.3.6-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.6-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.6-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.6-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	3.3.6-amzn-1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.6-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.6-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.6-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.6-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.6-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.3.6-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	2.4.17-amzn-3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-3	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-3	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-3	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-8	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-8	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-8	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-8	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-8	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	3.1.3-amzn-8	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-8	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.14.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.14.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.14.0-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.14.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.4.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter

Componente	Versión	Descripción
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.3	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.283-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.283-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.283-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	426-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	426-amzn-0	Service para ejecutar partes de una consulta.
trino-client	426-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.4.1-amzn-2	Clientes de línea de comando de Spark.
spark-history-server	3.4.1-amzn-2	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.4.1-amzn-2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.4.1-amzn-2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.08.1-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.10.2-amzn-6	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-6	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.15.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.15.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.15.0

Registro de cambios de la versión 6.15.0 y notas de la versión

Date	Evento	Descripción
17-11-2020	Publicación de documentos	Primera publicación de las notas de la versión 6.15.0 de Amazon EMR
2023-11-17	Implementación completa	Amazon EMR 6.15.0 se ha implementado completamente en todas las regiones compatibles
2023-11-13	Versión inicial	Implementación inicial de Amazon EMR 6.15.0 en regiones comerciales limitadas

Amazon EMR, versión 6.14.0

Versiones de las aplicaciones de la versión 6.14.0

Esta versión incluye las siguientes aplicaciones: Delta,,, Flink, Ganglia, HBase,, HCatalog, Hadoop, Hive, Hudi,, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, PhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
AWS SDK for Java	1.12.543	1.12,513	1.12.490	1.12.446
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.2.0
Flink	1.17.1	1.17.0	1.17.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
HBase	2.4.17	2.4.17	2.4.17	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,281	0,281	0,279
Spark	3.4.1	3.4.1	3.4.0	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
Trino (PrestoSQL)	422	414	414	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.14.0

Las siguientes notas de la versión incluyen información sobre la versión 6.14.0 de Amazon EMR. Los cambios son respecto a la versión 6.13.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.14.0.

Nuevas características

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- El Escalado administrado de Amazon EMR ya está disponible en la región Asia-Pacífico (Yakarta) `ap-southeast-3` para los clústeres que cree con Amazon EMR 6.14.0 y versiones posteriores.

Cambios, mejoras y problemas resueltos

- La versión 6.14.0 optimiza la administración de registros con Amazon EMR que se ejecuta en Amazon. EC2 Como resultado, es posible que observe una pequeña reducción en los costos de almacenamiento de los registros de su clúster.
- La versión 6.14.0 mejora el flujo de trabajo de escalado para tener en cuenta las diferentes instancias principales que tienen una variación sustancial de tamaño en sus volúmenes de Amazon EBS. Esta mejora se aplica únicamente a los nodos principales; las operaciones de reducción vertical de la escala de los nodos de tareas no se ven afectadas.
- La versión 6.14.0 mejora la forma en que Amazon EMR interactúa con aplicaciones de código abierto, como Apache Hadoop YARN ResourceManager and HDFS NameNode. Esta mejora reduce el riesgo de que se produzcan retrasos operativos relacionados con el escalado de clústeres y mitiga las fallas en el arranque que se producen debido a problemas de conectividad con las aplicaciones de código abierto.

- La versión 6.14.0 optimiza la instalación de las aplicaciones durante el lanzamiento del clúster. Esto reduce la cantidad necesaria de tiempo para que arranque el clúster para determinadas combinaciones de aplicaciones de Amazon EMR.
- La versión 6.14.0 corrige un problema por el que las operaciones de reducción vertical de la escala del clúster se interrumpen cuando un clúster que se ejecuta en una VPC con un dominio personalizado experimenta el reinicio de un nodo de tareas o nodo principal.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich), Asia Pacífico (Malasia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	17 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	11 de septiembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

Versiones predeterminadas de Java para la versión 6.14.0

Las versiones 6.12.0 y posteriores de Amazon EMR son compatibles con todas las aplicaciones con Amazon Corretto 8 de forma predeterminada, excepto Trino. En el caso de Trino, Amazon EMR es compatible con Amazon Corretto 17 de forma predeterminada a partir de la versión 6.9.0 de Amazon EMR. Amazon EMR también admite algunas aplicaciones con Amazon Corretto 11 y 17. Esas aplicaciones se indican en la siguiente tabla. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Si bien Amazon EMR es compatible con Amazon Corretto 11 y 17 en Apache Spark, Apache Hadoop y Apache Hive, es posible que el rendimiento de algunas cargas de trabajo disminuya cuando utilice estas versiones de Corretto. Se recomienda que pruebe sus cargas de trabajo antes de cambiar los valores predeterminados.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 6.14.0:

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
ZooKeeper	8

Versiones de los componentes de la versión 6.14.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la

comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.4.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
delta-standalone-connectors	0.6.0	Los conectores de Delta ofrecen diferentes tiempos de ejecución para integrar Delta Lake con motores como Flink, Hive y Presto.
emr-ddb	5.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.7.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.11.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter

Componente	Versión	Descripción
emr-s3-dist-cp	2.28.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.7.0	Conector S3Select de EMR
emr-wal-cli	1.1.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.59.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.17.1-amzn-0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.17.1-amzn-0	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	3.3.3-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-6	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-6	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	3.3.3-amzn-6	Servicio de YARN para la administración de contenedores en un nodo individual.
hadoop-yarn-resourcemanager	3.3.3-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.17-amzn-2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-2	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-2	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-2	Herramienta de reparación para HBase clústeres de Apache.

Componente	Versión	Descripción
hcatalog-client	3.1.3-amzn-7	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-7	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-7	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-7	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-7	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-7	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-7	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.13.1-amzn-2	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.13.1-amzn-2	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hudi-trino	0.13.1-amzn-2	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.13.1-amzn-2	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.3.1-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.2.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.3	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.281-amzn-2	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.281-amzn-2	Service para ejecutar partes de una consulta.
presto-client	0.281-amzn-2	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	422-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	422-amzn-0	Service para ejecutar partes de una consulta.
trino-client	422-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.4.1-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	3.4.1-amzn-1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.4.1-amzn-1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.4.1-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.06.0-amzn-2	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-5	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-5	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.14.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.14.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Clasificaciones	Descripción	Acciones de reconfiguración
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.14.0

Registro de cambios de la versión 6.14.0 y notas de la versión

Date	Evento	Descripción
2023-11-02	Implementación completa	Amazon EMR 6.14.0 se ha implementado completamente en todas las regiones compatibles
2023-10-10	Publicación de documentos	Primera publicación de las notas de la versión 6.14.0 de Amazon EMR
2023-10-04	Versión inicial	Implementación inicial de Amazon EMR 6.14.0 en regiones comerciales limitadas

Amazon EMR, versión 6.13.0

Versiones de las aplicaciones de la versión 6.13.0

Esta versión incluye las siguientes aplicaciones: Delta,,, Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue,, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNet, OoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
AWS SDK for Java	1.12.513	1.12.490	1.12.446	1.12.446
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.2.0	2.2.0
Flink	1.17.0	1.17.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
HBase	2.4.17	2.4.17	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,281	0,279	0,279
Spark	3.4.1	3.4.0	3.3.2	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
Trino (PrestoSQL)	414	414	410	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.13.0

Las siguientes notas de la versión incluyen información sobre la versión 6.13.0 de Amazon EMR. Los cambios son respecto a la versión 6.12.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.13.0.

Nuevas características

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

Cambios, mejoras y problemas resueltos

- La versión 6.13.0 mejora el daemon de administración de registros de Amazon EMR para garantizar que todos los registros se carguen a una cadencia normal en Amazon S3 cuando se emite un comando de terminación de clúster. Esto facilita que los clústeres se terminen de forma más rápida.
- La versión 6.13.0 mejora las capacidades de administración de registros de Amazon EMR para garantizar la carga coherente y puntual de todos los archivos de registro a Amazon S3. Esto beneficia especialmente a los clústeres de EMR de ejecución prolongada.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

Versiones predeterminadas de Java para la versión 6.13.0

Las versiones 6.12.0 y posteriores de Amazon EMR son compatibles con todas las aplicaciones con Amazon Corretto 8 de forma predeterminada, excepto Trino. En el caso de Trino, Amazon EMR es compatible con Amazon Corretto 17 de forma predeterminada a partir de la versión 6.9.0 de Amazon EMR. Amazon EMR también admite algunas aplicaciones con Amazon Corretto 11 y 17. Esas aplicaciones se indican en la siguiente tabla. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Si bien Amazon EMR es compatible con Amazon Corretto 11 y 17 en Apache Spark, Apache Hadoop y Apache Hive, es posible que el rendimiento de algunas cargas de trabajo disminuya cuando utilice estas versiones de Corretto. Se recomienda que pruebe sus cargas de trabajo antes de cambiar los valores predeterminados.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 6.13.0:

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
ZooKeeper	8

Versiones de los componentes de la versión 6.13.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la

comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.4.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
delta-standalone-connectors	0.6.0	Los conectores de Delta ofrecen diferentes tiempos de ejecución para integrar Delta Lake con motores como Flink, Hive y Presto.
emr-ddb	5.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.6.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.10.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter

Componente	Versión	Descripción
emr-s3-dist-cp	2.27.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.6.0	Conector S3Select de EMR
emr-wal-cli	1.1.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.58.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.17.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.17.0	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	3.3.3-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-5	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	3.3.3-amzn-5	Servicio de YARN para la administración de contenedores en un nodo individual.
hadoop-yarn-resourcemanager	3.3.3-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.17-amzn-1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-1	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-1	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-1	Herramienta de reparación para HBase clústeres de Apache.

Componente	Versión	Descripción
hcatalog-client	3.1.3-amzn-6	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-6	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-6	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-6	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-6	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-6	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-6	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.13.1-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.13.1-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hudi-trino	0.13.1-amzn-1	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.13.1-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.3.0-amzn-1	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.2.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.3	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.281-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.281-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.281-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	414-amzn-1	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	414-amzn-1	Service para ejecutar partes de una consulta.
trino-client	414-amzn-1	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.4.1-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.4.1-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.4.1-amzn-0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.4.1-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.06.0-amzn-1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-4	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-4	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.13.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.13.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Clasificaciones	Descripción	Acciones de reconfiguración
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.13.0

Registro de cambios de la versión 6.13.0 y notas de la versión

Date	Evento	Descripción
2023-09-23	Implementación completa	Amazon EMR 6.13.0 se ha implementado completamente en todas las regiones compatibles
2023-09-12	Publicación de documentos	Las notas de la versión 6.13.0 de Amazon EMR se publicaron por primera vez
2023-09-01	Versión inicial	Implementación inicial de Amazon EMR 6.13.0 en regiones comerciales limitadas

Amazon EMR, versión 6.12.0

Versiones de las aplicaciones de la versión 6.12.0

Esta versión incluye las siguientes aplicaciones: Delta,,, Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue,, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNet, OoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
AWS SDK for Java	1.12.490	1.12.446	1.12.446	1.12.397
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.2.0	2.2.0	2.2.0
Flink	1.17.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
HBase	2.4.17	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.10.0
Iceberg	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,281	0,279	0,279	0,278
Spark	3.4.0	3.3.2	3.3.2	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
Trino (PrestoSQL)	414	410	410	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.12.0

Las siguientes notas de la versión incluyen información sobre la versión 6.12.0 de Amazon EMR. Los cambios son respecto a la versión 6.11.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.12.0.

Nuevas características

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Las versiones 6.12.0 y posteriores de Amazon EMR admiten la integración de LDAP con Apache Livy, Apache Hive through HiveServer 2 (HS2), Trino, Presto y Hue. También puede instalar Apache Spark y Apache Hadoop en un clúster de EMR que utilice la versión 6.12.0 o posteriores y configurarlos para que utilicen LDAP. Para más información, consulte Uso de servidores de Active Directory o LDAP para la autenticación con Amazon EMR.

Cambios, mejoras y problemas resueltos

- La versión 6.12.0 y posteriores de Amazon EMR proporcionan compatibilidad con el tiempo de ejecución de Java 11 para Flink. Para obtener más información, consulte Configuración de Flink para que se ejecute con Java 11.
- La versión 6.12.0 agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se detenga indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.

- La versión 6.12.0 corrige un problema que provocaba que las operaciones de reducción vertical de los clústeres se interrumpieran cuando un nodo principal que se encontraba en proceso de retirada estable pasara a un estado incorrecto por cualquier motivo antes de retirarse por completo.
- La versión 6.12.0 mejora la lógica de reducción vertical de los clústeres para que el clúster no intente reducir verticalmente los nodos principales por debajo de la configuración de la configuración del factor de replicación de HDFS del clúster. Esto se ajusta a sus requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- La versión 6.12.0 mejora el rendimiento y la eficiencia del servicio de supervisión de estado de Amazon EMR al aumentar la velocidad a la que registra los cambios de estado de las instancias. Esta mejora reduce la posibilidad de que se degrade el rendimiento de los nodos del clúster que ejecutan varias herramientas de cliente personalizadas o aplicaciones de terceros.
- La versión 6.12.0 mejora el rendimiento del daemon de administración de registros en el clúster para Amazon EMR. Como resultado, hay menos probabilidades de que se degrade el rendimiento con los clústeres de EMR que ejecutan pasos con una alta simultaneidad.
- Con la versión 6.12.0 de Amazon EMR, el daemon de administración de registros se ha actualizado para identificar todos los registros que están en uso activo con controladores de archivos abiertos en el almacenamiento de instancias local y los procesos asociados. Esta actualización garantiza que Amazon EMR elimine correctamente los archivos y recupere espacio de almacenamiento una vez archivados los registros en Amazon S3.
- La versión 6.12.0 incluye una mejora del daemon de administración de registros que elimina los directorios escalonados vacíos y no utilizados del sistema de archivos del clúster local. Un número excesivamente elevado de directorios vacíos puede reducir el rendimiento de los daemons de Amazon EMR y provocar una sobreutilización del disco.
- La versión 6.12.0 permite la rotación de registros para los registros del servidor YARN Timeline. Esto minimiza los escenarios de sobreutilización del disco, especialmente en el caso de clústeres de ejecución prolongada.
- El tamaño del volumen raíz predeterminado ha aumentado a 15 GB en la versión 6.10.0 y posteriores de Amazon EMR. Las versiones anteriores tienen un tamaño de volumen raíz predeterminado de 10 GB.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Oregón) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

Versiones predeterminadas de Java para la versión 6.12.0

Las versiones 6.12.0 y posteriores de Amazon EMR son compatibles con todas las aplicaciones con Amazon Corretto 8 de forma predeterminada, excepto Trino. En el caso de Trino, Amazon EMR es compatible con Amazon Corretto 17 de forma predeterminada a partir de la versión 6.9.0 de Amazon EMR. Amazon EMR también admite algunas aplicaciones con Amazon Corretto 11 y 17. Esas aplicaciones se indican en la siguiente tabla. Si desea cambiar la JVM predeterminada del clúster, siga las instrucciones de Configuración de aplicaciones para utilizar una máquina virtual de Java específica para cada aplicación que se ejecute en el clúster. Solo se puede usar una versión de tiempo de ejecución de Java en un clúster. Amazon EMR no admite la ejecución de distintos nodos o aplicaciones en distintas versiones de tiempo de ejecución en el mismo clúster.

Si bien Amazon EMR es compatible con Amazon Corretto 11 y 17 en Apache Spark, Apache Hadoop y Apache Hive, es posible que el rendimiento de algunas cargas de trabajo disminuya cuando utilice estas versiones de Corretto. Se recomienda que pruebe sus cargas de trabajo antes de cambiar los valores predeterminados.

En la siguiente tabla, se muestran las versiones predeterminadas de Java para las aplicaciones en Amazon EMR 6.12.0:

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Delta	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8

Aplicación	Versión de Java o Amazon Corretto (el valor predeterminado está en negrita)
Livy	17, 11, 8
Oozie	17, 11, 8
Phoenix	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
ZooKeeper	8

Versiones de los componentes de la versión 6.12.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la

comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.4.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
delta-standalone-connectors	0.6.0	Los conectores de Delta ofrecen diferentes tiempos de ejecución para integrar Delta Lake con motores como Flink, Hive y Presto.
emr-ddb	5.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.5.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.9.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter

Componente	Versión	Descripción
emr-s3-dist-cp	2.26.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.5.0	Conector S3Select de EMR
emr-wal-cli	1.1.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.57.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.17.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.17.0	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	3.3.3-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	3.3.3-amzn-4	Servicio de YARN para la administración de contenedores en un nodo individual.
hadoop-yarn-resourcemanager	3.3.3-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.17-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.17-amzn-0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.17-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.4.17-amzn-0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.17-amzn-0	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.17-amzn-0	Herramienta de reparación para HBase clústeres de Apache.

Componente	Versión	Descripción
hcatalog-client	3.1.3-amzn-5	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-5	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-5	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-5	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-5	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-5	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-5	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.13.1-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.13.1-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hudi-trino	0.13.1-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.13.1-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.3.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.2.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	4.7.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.3	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.3	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.281-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.281-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.281-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	414-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	414-amzn-0	Service para ejecutar partes de una consulta.
trino-client	414-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.4.0-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.4.0-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.4.0-amzn-0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.4.0-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.06.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-3	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-3	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.12.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.12.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Clasificaciones	Descripción	Acciones de reconfiguración
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Cambie la página de inicio de Java de KMS de Hadoop	Not available.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.12.0

Registro de cambios de la versión 6.12.0 y notas de la versión

Date	Evento	Descripción
2023-07-27	Documentación actualizada	Actualice las opciones de Java para la versión 6.12 y agregue el tutorial de Oozie para actualizar la JVM
2023-07-21	Implementación completa	Amazon EMR 6.12.0 se ha implementado completamente en todas las regiones compatibles
2023-07-21	Publicación de documentos	Las notas de la versión 6.12.0 de Amazon EMR se publicaron por primera vez
2023-07-12	Versión inicial	Implementación inicial de Amazon EMR 6.12.0 en regiones comerciales limitadas

Amazon EMR, versión 6.11.1

Versiones de las aplicaciones de la versión 6.11.1

Esta versión incluye las siguientes aplicaciones: Delta,,, Flink, Ganglia, HBase, HCatalog, Hadoop,, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, MXNet, OoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
AWS SDK for Java	1.12.446	1.12.446	1.12.397	1.12.397
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.2.0
Flink	1.16.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
HBase	2.4.15	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.5.0	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,279	0,279	0,278	0,278
Spark	3.3.2	3.3.2	3.3.1	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
Trino (PrestoSQL)	410	410	403	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.11.1

Las siguientes notas de la versión incluyen información sobre la versión 6.11.1 de Amazon EMR. Los cambios son respecto a la versión 6.11.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.11.1.

Cambios, mejoras y problemas resueltos

- Debido a la contención de bloqueo, un nodo puede entrar en un interbloqueo si se agrega o elimina al mismo tiempo que intenta retirarse. Como resultado, el administrador de recursos de Hadoop (YARN) deja de responder y afecta a todos los contenedores entrantes y en ejecución actualmente.
- Esta versión incluye un cambio que permite que los clústeres de alta disponibilidad se recuperen de un estado de error tras el reinicio.
- Esta versión incluye correcciones de seguridad para Hue y HBase.
- Esta versión corrige un problema por el que los clústeres que ejecutan cargas de trabajo en Spark con Amazon EMR podrían recibir silenciosamente resultados incorrectos con `contains`, `startsWith`, `endsWith` y `like`. Este problema se produce cuando utiliza las expresiones en campos particionados que tienen metadatos en Hive3 Metastore Server (HMS) de Amazon EMR.
- Esta versión corrige un problema con la limitación en Glue cuando no hay funciones definidas por el usuario (UDF).
- Esta versión corrige un problema que elimina los registros de contenedores mediante el servicio de agregación de registros de nodos antes de que el insertador de registros pudiera enviarlos a S3 en caso de retirada de YARN.
- Esta versión corrige un problema con las métricas del FairShare programador cuando Node Label está habilitado para Hadoop.

- Esta versión corrige un problema que afectaba al rendimiento de Spark cuando se establecía un valor `true` predeterminado para la configuración `spark.yarn.heterogeneousExecutors.enabled` en `spark-defaults.conf`.
- Esta versión corrige un problema que provocaba que la tarea de reducción no pudiera leer los datos de mezclas aleatorias. El problema provocaba errores en las consultas de Hive debido a un error de memoria dañada.
- Esta versión agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se detenga indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.
- Esta versión mejora la lógica de reducción vertical del clúster para que el clúster no intente reducir verticalmente los nodos principales por debajo de la configuración del factor de replicación de HDFS del clúster. Esto se ajusta a sus requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- El daemon de administración de registros se ha actualizado para identificar todos los registros que están en uso activo con controladores de archivos abiertos en el almacenamiento de instancias local y los procesos asociados. Esta actualización garantiza que Amazon EMR elimine correctamente los archivos y recupere espacio de almacenamiento una vez archivados los registros en Amazon S3.
- Esta versión incluye una mejora del daemon de administración de registros que elimina los directorios escalonados vacíos y no utilizados del sistema de archivos del clúster local. Un número excesivamente elevado de directorios vacíos puede reducir el rendimiento de los daemons de Amazon EMR y provocar una sobreutilización del disco.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich), Asia Pacífico (Malasia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

Versiones de los componentes de la versión 6.11.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.2.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
delta-standalone-connectors	0.6.0	Los conectores de Delta ofrecen diferentes tiempos de ejecución para integrar Delta Lake con motores como Flink, Hive y Presto.
emr-ddb	5.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.8.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.25.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.4.0	Conector S3Select de EMR
emr-wal-cli	1.1.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.56.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.16.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.16.0	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustad o para aplicaciones del ecosistema de Hadoop junto con el agente de monitoriz ación de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.3.3-amzn-3.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-3.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-3.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-3.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-3.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-3.1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	3.3.3-amzn-3.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-3.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.3-amzn-3.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-3.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-3.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.15-amzn-1.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.15-amzn-1.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.15-amzn-1.1	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	2.4.15-amzn-1.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.15-amzn-1.1	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.15-amzn-1.1	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-4.1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-4.1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-4.1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-4.1	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-4.1	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-4.1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server2	3.1.3-amzn-4.1	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.13.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.13.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.13.0-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.13.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.2.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso

Componente	Versión	Descripción
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.2	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.279-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.279-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.279-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	410-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	410-amzn-0	Service para ejecutar partes de una consulta.
trino-client	410-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.2-amzn-0.1	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	3.3.2-amzn-0.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.2-amzn-0.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.3.2-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.02.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-2.1	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-2.1	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.11.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.11.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.11.1

Registro de cambios de la versión 6.11.1 y notas de la versión

Date	Evento	Descripción
30-08-2020	Actualización de las notas de la versión	Se han agregado varias correcciones relacionadas con el plano de control a las notas de la versión
2023-08-21	Publicación de documentos	Las notas de la versión 6.11.1 de Amazon EMR se publicaron por primera vez

Date	Evento	Descripción
2023-08-16	Implementación completa	Amazon EMR 6.11.1 se ha implementado completamente en todas las regiones compatibles
2023-08-04	Versión inicial	Amazon EMR 6.11.1 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 6.11.0

Versiones de las aplicaciones de la versión 6.11.0

Esta versión incluye las siguientes aplicaciones: Delta,,,, Flink, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi,, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, MXNetOoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
AWS SDK for Java	1.12.446	1.12.397	1.12.397	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.1.0
Flink	1.16.0	1.16.0	1.16.0	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.13
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0
Hue	4.11.0	4.10.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,279	0,278	0,278	0,276
Spark	3.3.2	3.3.1	3.3.1	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	410	403	403	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.11.0

Las siguientes notas de la versión incluyen información sobre la versión 6.11.0 de Amazon EMR. Los cambios son respecto a la versión 6.10.0. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Nuevas características

- Amazon EMR 6.11.0 es compatible con Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0 y PrestoDB 0.279-amzn-0.

Cambios, mejoras y problemas resueltos

- Con Amazon EMR 6.11.0, el conector de DynamoDB se ha actualizado a la versión 5.0.0. La versión 5.0.0 usa. AWS SDK for Java 2.x En las versiones anteriores se utilizaba la versión AWS SDK para Java 1.x. Como resultado de esta actualización, se recomienda encarecidamente que pruebe el código antes de utilizar el conector de DynamoDB con Amazon EMR 6.11.
- Cuando el conector de DynamoDB para Amazon EMR 6.11.0 llama al servicio DynamoDB, utiliza el valor de región que usted proporciona para la propiedad `dynamodb.endpoint`. Se recomienda que también configure `dynamodb.region` cuando utilice `dynamodb.endpoint` y que ambas propiedades tengan la misma Región de AWS. Si lo usa `dynamodb.endpoint` y no lo configura`dynamodb.region`, el conector de DynamoDB para Amazon EMR 6.11.0 devolverá una excepción de región no válida e intentará conciliar la Región de AWS información del servicio de metadatos de instancias de Amazon (IMDS). EC2 Si el conector no puede recuperar la región del IMDS, el valor predeterminado será Este de EE. UU. (Norte de Virginia) (`us-east-1`). El siguiente error es un ejemplo de la excepción de región no válida que puede producirse si no configura la `dynamodb.region` propiedad correctamente: `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` para obtener más información sobre las clases afectadas por la AWS SDK para Java actualización a la versión 2.x, consulte la confirmación de actualización AWS SDK para Java de la versión 1.x a la 2.x (#175) en el GitHub repositorio del conector Amazon EMR - DynamoDB.
- Esta versión corrige un problema por el que los datos de las columnas se convierten en `NULL` cuando se utiliza Delta Lake para almacenar los datos de la tabla Delta en Amazon S3 tras una operación de cambio de nombre de columna. Para más información sobre esta característica experimental de Delta Lake, consulte Operación de cambio de nombre de columnas en la Guía del usuario de Delta Lake.
- La versión 6.11.0 corrige un problema que podía producirse al crear un nodo de periferia mediante la replicación de uno de los nodos principales de un clúster con varios nodos principales. El nodo de periferia replicado podría provocar retrasos en las operaciones de reducción vertical o provocar un uso elevado de la memoria en los nodos principales. Para obtener más información sobre cómo crear un nodo perimetral para comunicarse con su clúster de EMR, consulte Edge Node Creator en el `aws-samples` repositorio de. GitHub
- La versión 6.11.0 mejora el proceso de automatización que Amazon EMR utiliza para volver a montar los volúmenes de Amazon EBS en una instancia tras un reinicio.
- La versión 6.11.0 corrige un problema que provocaba brechas intermitentes en las métricas de Hadoop que Amazon EMR publica en Amazon. CloudWatch

- La versión 6.11.0 corrige un problema con los clústeres de EMR por el que se interrumpe una actualización del archivo de configuración de YARN que contiene la lista de nodos excluidos del clúster debido a una sobreutilización del disco. La actualización incompleta dificulta las futuras operaciones de reducción vertical de los clústeres. Esta versión garantiza que el clúster se mantenga en buen estado y que las operaciones de escalado funcionen según lo previsto.
- El tamaño del volumen raíz predeterminado ha aumentado a 15 GB en la versión 6.10.0 y posteriores de Amazon EMR. Las versiones anteriores tienen un tamaño de volumen raíz predeterminado de 10 GB.
- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. Este enfoque también puede provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` se establece como `false` en `yarn-site.xml` para resolver este problema.

  Si bien la corrección soluciona los problemas que ha introducido YARN-9608, es posible que los trabajos de Hive presenten errores debido a la pérdida de datos de mezclas aleatorias en los clústeres que tienen habilitado el escalado administrado. Se ha mitigado ese riesgo en esta versión al configurar también `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para las cargas de trabajo de Hive. Esta configuración solo está disponible con la versión 6.11.0 y posteriores de Amazon EMR.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-`

`labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

Versiones de los componentes de la versión 6.11.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.2.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
delta-standalone-connectors	0.6.0	Los conectores de Delta ofrecen diferentes tiempos de ejecución para integrar Delta Lake con motores como Flink, Hive y Presto.
emr-ddb	5.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	3.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.8.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.25.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.4.0	Conector S3Select de EMR
emr-wal-cli	1.1.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.56.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.16.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.16.0	Administración de recursos en nodos EMR para Apache Flink. JobManager

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.3.3-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	3.3.3-amzn-3	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.3-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.15-amzn-1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	2.4.15-amzn-1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.15-amzn-1	HBase cliente de línea de comandos.
hbase-rest-server	2.4.15-amzn-1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.15-amzn-1	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.15-amzn-1	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-4	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-4	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	3.1.3-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.13.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.13.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.13.0-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.13.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.2.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter

Componente	Versión	Descripción
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.2	Conectores de Apache Phoenix para Spark-3

Componente	Versión	Descripción
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.279-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.279-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.279-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	410-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	410-amzn-0	Service para ejecutar partes de una consulta.
trino-client	410-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.2-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.3.2-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.2-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.3.2-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	23.02.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.10.2-amzn-2	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-2	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.11.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.11.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.11.0

Registro de cambios de la versión 6.11.0 y notas de la versión

Date	Evento	Descripción
2023-08-21	Actualización	Se ha corregido un problema introducido con Hadoop 3.3.3.
2023-07-26	Actualización	Nuevas etiquetas de lanzamiento del sistema operativo `2.0.20230612.0` y `2.0.20230628.0` .
2023-06-09	Implementación completa	Amazon EMR 6.11.0 se ha implementado completam ente en todas las regiones compatibles
2023-06-09	Publicación de documentos	Las notas de la versión 6.11.0 de Amazon EMR se publicaro n por primera vez
2023-06-08	Versión inicial	Implementación inicial de Amazon EMR 6.11.0 en regiones comerciales limitadas

Amazon EMR, versión 6.10.1

Versiones de las aplicaciones de la versión 6.10.1

Esta versión incluye las siguientes aplicaciones: Delta,,,, Flink, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi,, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, MXNetOoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
AWS SDK for Java	1.12.397	1.12.397	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.1.0	2.1.0
Flink	1.16.0	1.16.0	1.15.2	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.13	2.4.13
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
Iceberg	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,278	0,278	0,276	0,276
Spark	3.3.1	3.3.1	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.10.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	403	403	398	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.10.1

Las siguientes notas de la versión incluyen información sobre la versión 6.10.1 de Amazon EMR. Los cambios son respecto a la versión 6.10.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.10.1.

Cambios, mejoras y problemas resueltos

- Debido a la contención de bloqueo, un nodo puede entrar en un interbloqueo si se agrega o elimina al mismo tiempo que intenta retirarse. Como resultado, el administrador de recursos de Hadoop (YARN) deja de responder y afecta a todos los contenedores entrantes y en ejecución actualmente.
- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. Este enfoque también puede provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` se establece como `false` en `yarn-site.xml` para resolver este problema.

  Si bien la corrección soluciona los problemas que ha introducido YARN-9608, es posible que los trabajos de Hive presenten errores debido a la pérdida de datos de mezclas aleatorias en los clústeres que tienen habilitado el escalado administrado. Se ha mitigado ese riesgo en esta versión al configurar también `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para las cargas de trabajo de Hive. Esta configuración solo está disponible con la versión 6.11.0 y posteriores de Amazon EMR.
- El recopilador de métricas no enviará ninguna métrica al plano de control después de una conmutación por error del nodo principal en los clústeres con la configuración de grupos de instancias.
- Esta versión incluye un cambio que permite que los clústeres de alta disponibilidad se recuperen de un estado de error tras el reinicio.
- Esta versión incluye correcciones de seguridad para Hue y HBase.
- Esta versión corrige un problema por el que los clústeres que ejecutan cargas de trabajo en Spark con Amazon EMR podrían recibir silenciosamente resultados incorrectos con `contains`,

`startsWith`, `endsWith` y `like`. Este problema se produce cuando utiliza las expresiones en campos particionados que tienen metadatos en Hive3 Metastore Server (HMS) de Amazon EMR.

- Esta versión corrige un problema con la limitación en Glue cuando no hay funciones definidas por el usuario (UDF).
- Esta versión corrige un problema que elimina los registros de contenedores mediante el servicio de agregación de registros de nodos antes de que el insertador de registros pudiera enviarlos a S3 en caso de retirada de YARN.
- Esta versión corrige un problema con las métricas del FairShare programador cuando Node Label está habilitado para Hadoop.
- Esta versión corrige un problema que afectaba al rendimiento de Spark cuando se establecía un valor `true` predeterminado para la configuración `spark.yarn.heterogeneousExecutors.enabled` en `spark-defaults.conf`.
- Esta versión corrige un problema que provocaba que la tarea de reducción no pudiera leer los datos de mezclas aleatorias. El problema provocaba errores en las consultas de Hive debido a un error de memoria dañada.
- Esta versión agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se detenga indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.
- Esta versión mejora la lógica de reducción vertical del clúster para que el clúster no intente reducir verticalmente los nodos principales por debajo de la configuración del factor de replicación de HDFS del clúster. Esto se ajusta a sus requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- El daemon de administración de registros se ha actualizado para identificar todos los registros que están en uso activo con controladores de archivos abiertos en el almacenamiento de instancias local y los procesos asociados. Esta actualización garantiza que Amazon EMR elimine correctamente los archivos y recupere espacio de almacenamiento una vez archivados los registros en Amazon S3.
- Esta versión incluye una mejora del daemon de administración de registros que elimina los directorios escalonados vacíos y no utilizados del sistema de archivos del clúster local. Un número excesivamente elevado de directorios vacíos puede reducir el rendimiento de los daemons de Amazon EMR y provocar una sobreutilización del disco.

- Esta versión corrige un problema que podía producirse al crear un nodo de periferia cuando se replicaba uno de los nodos principales de un clúster con varios nodos principales. El nodo de periferia replicado podría provocar retrasos en las operaciones de reducción vertical o provocar un uso elevado de la memoria en los nodos principales. Para obtener más información sobre cómo crear un nodo perimetral para comunicarse con su clúster de EMR, consulte Edge Node Creator en el `aws-samples` repositorio de. GitHub
- Esta versión mejora el proceso de automatización que Amazon EMR utiliza para volver a montar los volúmenes de Amazon EBS en una instancia tras un reinicio.
- Esta versión corrige un problema que provocaba brechas intermitentes en las métricas de Hadoop que Amazon EMR publica en Amazon. CloudWatch
- Esta versión corrige un problema con los clústeres de EMR por el que se interrumpe una actualización del archivo de configuración de YARN que contiene la lista de nodos excluidos del clúster debido a una sobreutilización del disco. La actualización incompleta dificulta las futuras operaciones de reducción vertical de los clústeres. Esta versión garantiza que el clúster se mantenga en buen estado y que las operaciones de escalado funcionen según lo previsto.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

Versiones de los componentes de la versión 6.10.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.2.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.7.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter

Componente	Versión	Descripción
emr-s3-dist-cp	2.24.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.3.0	Conector S3Select de EMR
emr-wal-cli	1.0.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.55.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.16.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.16.0	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	3.3.3-amzn-2.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-2.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-2.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-2.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-2.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-2.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-2.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-2.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	3.3.3-amzn-2.1	Servicio de YARN para la administración de contenedores en un nodo individual.
hadoop-yarn-resourcemanager	3.3.3-amzn-2.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-2.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.15-amzn-0.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.15-amzn-0.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.15-amzn-0.1	HBase cliente de línea de comandos.
hbase-rest-server	2.4.15-amzn-0.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.15-amzn-0.1	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.15-amzn-0.1	Herramienta de reparación para HBase clústeres de Apache.

Componente	Versión	Descripción
hcatalog-client	3.1.3-amzn-3.1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-3.1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-3.1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-3.1	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-3.1	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-3.1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-3.1	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.12.2-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.12.2-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hudi-trino	0.12.2-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.12.2-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.1.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.2.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0-SNAPSHOT	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.278.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.278.1-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.278.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	403-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	403-amzn-0	Service para ejecutar partes de una consulta.
trino-client	403-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.1-amzn-0.1	Clientes de línea de comando de Spark.
spark-history-server	3.3.1-amzn-0.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.1-amzn-0.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.3.1-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.12.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-1.1	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-1.1	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.10.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.10.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Clasificaciones	Descripción	Acciones de reconfiguración
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.pro perties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.10.1

Registro de cambios de la versión 6.10.1 y notas de la versión

Date	Evento	Descripción
30-08-2020	Actualización de las notas de la versión	Se han agregado varias correcciones relacionadas con el plano de control a las notas de la versión
2023-08-21	Publicación de documentos	Las notas de la versión 6.10.1 de Amazon EMR se publicaron por primera vez
2023-08-16	Implementación completa	Amazon EMR 6.10.1 se ha implementado completamente en todas las regiones compatibles
2023-08-04	Versión inicial	Amazon EMR 6.10.1 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 6.10.0

Versiones de las aplicaciones de la versión 6.10.0

Esta versión incluye las siguientes aplicaciones: Delta,,,, Flink, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi,, Hue, Iceberg, JupyterEnterpriseGateway,, JupyterHub, Livy, MXNetOoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
AWS SDK for Java	1.12.397	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.1.0	2.1.0	-
Flink	1.16.0	1.15.2	1.15.2	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.13	2.4.13	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.1.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,278	0,276	0,276	0,273
Spark	3.3.1	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.10.0	2.10.0	2.9.1
Tez	0.10.2	0.10.2	0.10.2	0.9.2
Trino (PrestoSQL)	403	398	398	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.10.0

Las siguientes notas de la versión incluyen información sobre la versión 6.10.0 de Amazon EMR. Los cambios son respecto a la versión 6.9.0. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Nuevas características

- Amazon EMR 6.10.0 es compatible con Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403 y PrestoDB 0.278.1.
- Amazon EMR 6.10.0 incluye un conector Trino-Hudi nativo que proporciona acceso de lectura a los datos de las tablas de Hudi. Puede activar el conector con `trino-cli --catalog hudi` y configurarlo según sus requisitos con `trino-connector-hudi`. La integración nativa con Amazon EMR significa que ya no necesitará utilizar `trino-connector-hive` para consultar tablas de Hudi. Para obtener una lista de las configuraciones compatibles con el nuevo conector, consulte la página Hudi connector de la documentación de Trino.
- Las versiones 6.10.0 y posteriores de Amazon EMR admiten la integración de Apache Zeppelin con Apache Flink. Para obtener más información, consulte Trabajar con trabajos de Flink de Zeppelin en Amazon EMR.

Problemas conocidos

- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. Este enfoque también puede provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  Para solucionar este problema en Amazon EMR 6.10.0, puede establecer el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` en `yarn-site.xml`. En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, la configuración se establece en `false` de manera predeterminada para resolver este problema.

Cambios, mejoras y problemas resueltos

- Amazon EMR 6.10.0 elimina la dependencia en `minimal-json.jar` de la integración de Amazon Redshift para Apache Spark y agrega automáticamente los archivos JAR necesarios relacionados con Spark-Redshift a la ruta de clases de los ejecutores de Spark: `spark-redshift.jar`, `spark-avro.jar` y `RedshiftJDBC.jar`.

- La versión 6.10.0 mejora el daemon de administración de registros en el clúster para supervisar las carpetas de registro adicionales del clúster de EMR. Esta mejora minimiza los escenarios de sobreutilización del disco.
- La versión 6.10.0 reinicia automáticamente el daemon de administración de registros en el clúster cuando se detiene. Esta mejora reduce el riesgo de que los nodos parezcan estar en mal estado debido a la sobreutilización del disco.
- Amazon EMR 6.10.0 admite puntos de conexión regionales para la asignación de usuarios de EMRFS.
- El tamaño del volumen raíz predeterminado ha aumentado a 15 GB en la versión 6.10.0 y posteriores de Amazon EMR. Las versiones anteriores tienen un tamaño de volumen raíz predeterminado de 10 GB.
- La versión 6.10.0 corrige un problema que provocaba que los trabajos de Spark se detuvieran cuando todos los ejecutores de Spark restantes se encontraban en un host de retirada con el administrador de recursos de YARN.
- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.`*`1`*). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-`

`labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

Versiones de los componentes de la versión 6.10.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de

Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.2.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.7.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que

Componente	Versión	Descripción
		incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.24.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.3.0	Conector S3Select de EMR
emr-wal-cli	1.0.0	CLI utilizada para eliminar o enumerar WAL de ERM.
emrfs	2.55.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.16.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.16.0	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.3.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-2	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	3.3.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.3.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.3.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.15-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.15-amzn-0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.15-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.4.15-amzn-0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.15-amzn-0	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hbase-operator-tools	2.4.15-amzn-0	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-3	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-3	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-3	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-3	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-3	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-3	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-3	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hudi	0.12.2-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.12.2-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.12.2-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.12.2-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	1.1.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.5.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.

Componente	Versión	Descripción
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.8.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0-SNAPSHOT	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.278.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.278.1-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.278.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	403-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	403-amzn-0	Service para ejecutar partes de una consulta.
trino-client	403-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.1-amzn-0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	3.3.1-amzn-0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.1-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.3.1-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.12.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-1	La aplicación YARN de tez y bibliotecas.
tez-on-worker	0.10.2-amzn-1	La aplicación YARN de tez y bibliotecas para nodos de trabajo.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.10.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.10.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Cambiar los valores en el archivo hudi.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-logback	Cambie los valores del archivo kms-logback.xml de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.10.0

Registro de cambios de la versión 6.10.0 y notas de la versión

Date	Evento	Descripción
2023-08-21	Actualización	Se ha agregado un problema conocido introducido con Hadoop 3.3.3.
2023-07-26	Actualización	Nuevas etiquetas de lanzamiento del sistema operativo `2.0.20230612.0` y `2.0.20230628.0` .

Date	Evento	Descripción
2023-03-02	Implementación completa	Amazon EMR 6.10 se ha implementado completamente en todas las regiones compatibles
2023-03-02	Publicación de documentos	Las notas de la versión 6.10 de Amazon EMR se publicaron por primera vez
2023-02-27	Versión inicial	Amazon EMR 6.10 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 6.9.1

Versiones de las aplicaciones de la versión 6.9.1

Esta versión incluye las siguientes aplicaciones: Delta,,, Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue,, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNet, OoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	2.1.0	-	-
Flink	1.15.2	1.15.2	1.15.1	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.13	2.4.12	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,276	0,276	0,273	0,273
Spark	3.3.0	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.10.0	2.9.1	2.9.1
Tez	0.10.2	0.10.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	398	388	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

Notas de la versión 6.9.1

Las siguientes notas de la versión incluyen información sobre la versión 6.9.1 de Amazon EMR. Los cambios son respecto a la versión 6.9.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.9.1.

Cambios, mejoras y problemas resueltos

- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. Este enfoque también puede provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` se establece como `false` en `yarn-site.xml` para resolver este problema.

Si bien la corrección soluciona los problemas que ha introducido YARN-9608, es posible que los trabajos de Hive presenten errores debido a la pérdida de datos de mezclas aleatorias en los clústeres que tienen habilitado el escalado administrado. Se ha mitigado ese riesgo en esta versión al configurar también `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para las cargas de trabajo de Hive. Esta configuración solo está disponible con la versión 6.11.0 y posteriores de Amazon EMR.

- El recopilador de métricas no enviará ninguna métrica al plano de control después de una conmutación por error del nodo principal en los clústeres con la configuración de grupos de instancias.
- Esta versión elimina los reintentos de las solicitudes HTTP con errores a los puntos de conexión del recopilador de métricas.
- Esta versión incluye un cambio que permite que los clústeres de alta disponibilidad se recuperen de un estado de error tras el reinicio.
- Esta versión corrige un problema por el que las grandes excepciones creadas por los usuarios UIDs provocaban excepciones de desbordamiento.
- Esta versión corrige los problemas de tiempo de espera con el proceso de reconfiguración de Amazon EMR.
- Esta versión incluye correcciones de seguridad.
- Esta versión corrige un problema por el que los clústeres que ejecutan cargas de trabajo en Spark con Amazon EMR podrían recibir silenciosamente resultados incorrectos con `contains`, `startsWith`, `endsWith` y `like`. Este problema se produce cuando utiliza las expresiones en campos particionados que tienen metadatos en Hive3 Metastore Server (HMS) de Amazon EMR.
- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.

- Es posible que Hive pierda datos si utiliza HDFS como directorio temporal y ha habilitado la combinación de archivos pequeños, y la tabla contiene rutas de partición estáticas.
- Esta versión corrige un problema de rendimiento con Hive si la combinación de archivos pequeños (deshabilitada de forma predeterminada) se habilita al final del trabajo de ETL.
- Esta versión corrige un problema con la limitación en Glue cuando no hay funciones definidas por el usuario (UDF).
- Esta versión corrige un problema que elimina los registros de contenedores mediante el servicio de agregación de registros de nodos antes de que el insertador de registros pudiera enviarlos a S3 en caso de retirada de YARN.
- Esta versión corrige la gestión de los archivos compactados o archivados con un seguimiento persistente de los archivos de almacenamiento. HBase
- Esta versión corrige un problema que afectaba al rendimiento de Spark cuando se establecía un valor `true` predeterminado para la configuración `spark.yarn.heterogeneousExecutors.enabled` en `spark-defaults.conf`.
- Esta versión corrige un problema que provocaba que la tarea de reducción no pudiera leer los datos de mezclas aleatorias. El problema provocaba errores en las consultas de Hive debido a un error de memoria dañada.
- Esta versión corrige un problema que provocaba un error en el aprovisionador de nodos si el servicio HDFS NameNode (NN) se bloqueaba en modo seguro durante la sustitución del nodo.
- Esta versión agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se detenga indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.
- Esta versión mejora la lógica de reducción vertical del clúster para que el clúster no intente reducir verticalmente los nodos principales por debajo de la configuración del factor de replicación de HDFS del clúster. Esto se ajusta a sus requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- El daemon de administración de registros se ha actualizado para identificar todos los registros que están en uso activo con controladores de archivos abiertos en el almacenamiento de instancias local y los procesos asociados. Esta actualización garantiza que Amazon EMR elimine correctamente los archivos y recupere espacio de almacenamiento una vez archivados los registros en Amazon S3.

- Esta versión incluye una mejora del daemon de administración de registros que elimina los directorios escalonados vacíos y no utilizados del sistema de archivos del clúster local. Un número excesivamente elevado de directorios vacíos puede reducir el rendimiento de los daemons de Amazon EMR y provocar una sobreutilización del disco.
- Esta versión corrige un problema que podía producirse al crear un nodo de periferia cuando se replicaba uno de los nodos principales de un clúster con varios nodos principales. El nodo de periferia replicado podría provocar retrasos en las operaciones de reducción vertical o provocar un uso elevado de la memoria en los nodos principales. Para obtener más información sobre cómo crear un nodo perimetral para comunicarse con su clúster de EMR, consulte Edge Node Creator en el `aws-samples` repositorio de. GitHub
- Esta versión mejora el proceso de automatización que Amazon EMR utiliza para volver a montar los volúmenes de Amazon EBS en una instancia tras un reinicio.
- Esta versión corrige un problema que provocaba brechas intermitentes en las métricas de Hadoop que Amazon EMR publica en Amazon. CloudWatch
- Esta versión corrige un problema con los clústeres de EMR por el que se interrumpe una actualización del archivo de configuración de YARN que contiene la lista de nodos excluidos del clúster debido a una sobreutilización del disco. La actualización incompleta dificulta las futuras operaciones de reducción vertical de los clústeres. Esta versión garantiza que el clúster se mantenga en buen estado y que las operaciones de escalado funcionen según lo previsto.
- Esta versión mejora el daemon de administración de registros en el clúster para supervisar las carpetas de registro adicionales del clúster de EMR. Esta mejora minimiza los escenarios de sobreutilización del disco.
- Esta versión reinicia automáticamente el daemon de administración de registros en el clúster cuando se detiene. Esta mejora reduce el riesgo de que los nodos parezcan estar en mal estado debido a la sobreutilización del disco.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourn e), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central) , Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia), Canadá occidental (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourn e), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central) , Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia), Canadá occidental (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourn e), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central) , Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia), Canadá occidental (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourn e), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central) , Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia), Canadá occidental (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

Versiones de los componentes de la versión 6.9.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.1.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.6.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter

Componente	Versión	Descripción
emr-s3-dist-cp	2.23.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.2.0	Conector S3Select de EMR
emrfs	2.54,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.15.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.15.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	3.3.3-amzn-1.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-1.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-1.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-1.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-1.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-1.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-1.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-1.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	3.3.3-amzn-1.1	Servicio de YARN para la administración de contenedores en un nodo individual.
hadoop-yarn-resourcemanager	3.3.3-amzn-1.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-1.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.13-amzn-0.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.13-amzn-0.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.13-amzn-0.1	HBase cliente de línea de comandos.
hbase-rest-server	2.4.13-amzn-0.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.13-amzn-0.1	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.13-amzn-0.1	Herramienta de reparación para HBase clústeres de Apache.

Componente	Versión	Descripción
hcatalog-client	3.1.3-amzn-2.1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-2.1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-2.1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-2.1	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-2.1	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-2.1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-2.1	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.12.1-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.12.1-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hudi-trino	0.12.1-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.12.1-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	0.14.1-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.7.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.2.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0-SNAPSHOT	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.276-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.276-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.276-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	398-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	398-amzn-0	Service para ejecutar partes de una consulta.
trino-client	398-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.0-amzn-1.1	Clientes de línea de comando de Spark.
spark-history-server	3.3.0-amzn-1.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.0-amzn-1.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.3.0-amzn-1.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.08.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.10.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-0.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.9.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.9.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Clasificaciones	Descripción	Acciones de reconfiguración
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.9.1

Registro de cambios de la versión 6.9.1 y notas de la versión

Date	Evento	Descripción
30-08-2020	Actualización de las notas de la versión	Se han agregado varias correcciones relacionadas con el plano de control a las notas de la versión
2023-08-21	Publicación de documentos	Las notas de la versión 6.9.1 de Amazon EMR se publicaron por primera vez
2023-08-16	Implementación completa	Amazon EMR 6.9.1 se ha implementado completamente en todas las regiones compatibles
2023-08-04	Versión inicial	Amazon EMR 6.9.1 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 6.9.0

Versiones de las aplicaciones de la versión 6.9.0

Esta versión incluye las siguientes aplicaciones: Delta,,,, Flink, Ganglia,, HBase, HCatalog, Hadoop,, Hive, Hudi, Hue,, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNetOoziePhoenixPigPresto Spark, Sqoop, TensorFlow, TezTrino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	-	-	-
Flink	1.15.2	1.15.1	1.15.1	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.12	2.4.12	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,276	0,273	0,273	0,272
Spark	3.3.0	3.3.0	3.3.0	3.2.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.9.1	2.9.1	2.4.1
Tez	0.10.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	388	388	378
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.7

Notas de la versión 6.9.0

Las siguientes notas de la versión incluyen información sobre la versión 6.9.0 de Amazon EMR. Los cambios son respecto a la versión 6.8.0 de Amazon EMR. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Nuevas características

- La versión 6.9.0 de Amazon EMR es compatible con Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398 y Tez 0.10.2.
- La versión 6.9.0 de Amazon EMR incluye una nueva aplicación de código abierto, Delta Lake 2.1.0.
- La integración de Amazon Redshift para Apache Spark se incluye en las versiones 6.9.0 y posteriores de Amazon EMR. La integración nativa, que anteriormente era una herramienta de código abierto, es un conector de Spark que puede utilizar para crear aplicaciones de Apache Spark que leen y escriben datos en Amazon Redshift y Amazon Redshift sin servidor. Para obtener más información, consulte Uso de la integración de Amazon Redshift para Apache Spark con Amazon EMR .
- La versión 6.9.0 de Amazon EMR agrega compatibilidad para archivar registros en Amazon S3 durante la reducción vertical del clúster. Anteriormente, solo podía archivar archivos de registro en Amazon S3 durante la terminación del clúster. Esta nueva capacidad garantiza que los archivos de registro generados en el clúster persistan en Amazon S3 incluso después de terminar el nodo. Para más información, consulte Configurar el registro y la depuración de un clúster.
- Para admitir consultas de ejecución prolongada, Trino ahora incluye un mecanismo de ejecución tolerante a errores. La ejecución tolerante a errores mitiga los errores de las consultas al volver a intentar las consultas con errores o las tareas que las componen.
- Puede utilizar Apache Flink en Amazon EMR para el procesamiento unificado `BATCH` y `STREAM` de tablas de Apache Hive o metadatos de cualquier origen de tabla de Flink, como Iceberg, Kinesis o Kafka. Puedes especificar el catálogo de datos de AWS Glue como el metaalmacén de Flink mediante la API AWS Management Console AWS CLI, o Amazon EMR. Para obtener más información, consulte Configuración de Flink en Amazon EMR.
- Ahora puede especificar las funciones de tiempo de ejecución AWS Identity and Access Management (IAM) y el control de acceso AWS Lake Formation basado para las consultas de Apache Spark, Apache Hive y Presto en Amazon EMR en clústeres EC2 con Amazon AI Studio. SageMaker Para más información, consulte Configuración de roles en tiempo de ejecución para los pasos de Amazon EMR.

Problemas conocidos

- En el caso de la versión 6.9.0 de Amazon EMR, Trino no funciona en clústeres habilitados para Apache Ranger. Si tiene que utilizar Trino con Ranger, contacte con Soporte.

- Si utiliza la integración de Amazon Redshift para Apache Spark y tiene un valor de time, timetz, timestamp o timestamptz con una precisión de microsegundos en formato Parquet, el conector redondea los valores de tiempo al valor de milisegundos más cercano. Como solución alternativa, utilice el parámetro `unload_s3_format` de formato de descarga de texto.
- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Las conexiones a los clústeres de Amazon EMR desde Amazon SageMaker AI Studio pueden fallar de forma intermitente con un código de respuesta 403 Forbidden. Este error se produce cuando la configuración del rol de IAM en el clúster tarda más de 60 segundos. Como solución alternativa, puede instalar una revisión de Amazon EMR para permitir los reintentos y aumentar el tiempo de espera a un mínimo de 300 segundos. Siga los siguientes pasos para aplicar la acción de arranque al lanzar el clúster.

  1. Descargue el script de arranque y los archivos RPM del siguiente Amazon S3 URIs.

     ```
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
     ```

  2. Cargue los archivos del paso anterior a un bucket de Amazon S3 de su propiedad. El depósito debe estar en el mismo Región de AWS lugar en el que planea lanzar el clúster.

3. Incluya la siguiente acción de arranque cuando inicie el clúster de EMR. Sustituya *bootstrap_URI* y *RPM_URI* por la correspondiente URIs de Amazon S3.

```
--bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
```

- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- Apache Flink proporciona FileSystem conectores S3 FileSystem y Hadoop nativos, que permiten a las aplicaciones crear FileSink y escribir los datos en Amazon S3. Esto FileSink falla con una de las dos excepciones siguientes.

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
 supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError:
 org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;
```

```
                                        at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

Como solución alternativa, puede instalar una revisión de Amazon EMR, que soluciona el problema anterior en Flink. Siga los siguientes pasos para aplicar la acción de arranque al iniciar el clúster.

1. Descargue el archivo flink-rpm en su bucket de Amazon S3. La ruta de su RPM es `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`.
2. Descargue el script de arranque y los archivos RPM de Amazon S3 con los siguientes URI. *regionName*Reemplácelo por el Región de AWS lugar en el que planea lanzar el clúster.

   ```
   s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-
   rpms.sh
   ```

3. Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. En Amazon EMR 6.8.0 y 6.9.0, este enfoque también podría provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

   Con Amazon EMR 6.10.0, hay una solución alternativa para este problema: puede establecer el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` en `yarn-site.xml`. En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, la configuración se establece en `false` de manera predeterminada para resolver este problema.

Cambios, mejoras y problemas resueltos

- En el caso de la versión 6.9.0 y posteriores de Amazon EMR, todos los componentes instalados por Amazon EMR que utilizan bibliotecas Log4j utilizan la versión 2.17.1 o posteriores de Log4j.
- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. La versión 6.9.0 de Amazon EMR corrige este problema.

- Amazon EMR 6.9.0 agrega compatibilidad limitada para el control de acceso basado en Lake Formation con Apache Hudi al leer datos con Spark SQL. La compatibilidad es para consultas SELECT que utilizan Spark SQL y se limita al control de acceso en el nivel de columnas. Para más información, consulte Hudi y Lake Formation.
- Cuando utiliza Amazon EMR 6.9.0 para crear un clúster de Hadoop con etiquetas de nodo habilitadas, la API de métricas de YARN devuelve información agregada en todas las particiones, en lugar de la partición predeterminada. Para más información, consulte YARN-11414.
- Con la versión 6.9.0 de Amazon EMR, hemos actualizado Trino a la versión 398, que utiliza Java 17. La versión anterior compatible de Trino para Amazon EMR 6.8.0 era Trino 388 y se ejecutaba en Java 11. Para más información sobre este cambio, consulte las actualizaciones de Trino a Java 17 en el blog de Trino.
- Esta versión corrige un problema de desajuste de la secuencia de tiempo entre Apache BigTop y Amazon EMR EC2 en la secuencia de inicio del clúster. Este desajuste de la secuencia de tiempo se produce cuando un sistema intenta llevar a cabo dos o más operaciones al mismo tiempo en lugar de hacerlas en la secuencia correcta. Como resultado, algunas configuraciones de clústeres experimentaron tiempos de espera de inicio de instancias y tiempos de inicio de clústeres más lentos.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Europa (Zúrich)
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	12 de enero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

Versiones de los componentes de la versión 6.9.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
delta	2.1.0	Delta Lake es un formato de tabla abierto para conjuntos de datos analíticos de gran tamaño
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.6.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter

Componente	Versión	Descripción
emr-s3-dist-cp	2.23.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.2.0	Conector S3Select de EMR
emrfs	2.54,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.15.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.15.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	3.3.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.3.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.3.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.3.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.3.3-amzn-1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.3.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.3.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.3.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	3.3.3-amzn-1	Servicio de YARN para la administración de contenedores en un nodo individual.
hadoop-yarn-resourcemanager	3.3.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	3.3.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.13-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.13-amzn-0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.13-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.4.13-amzn-0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.13-amzn-0	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.13-amzn-0	Herramienta de reparación para HBase clústeres de Apache.

Componente	Versión	Descripción
hcatalog-client	3.1.3-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-2	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.12.1-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.12.1-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hudi-trino	0.12.1-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.12.1-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	0.14.1-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.7.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.2.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	6.0.0-SNAPSHOT	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	6.0.0	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.276-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.276-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.276-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	398-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	398-amzn-0	Service para ejecutar partes de una consulta.
trino-client	398-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.0-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	3.3.0-amzn-1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.0-amzn-1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.3.0-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.08.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.10.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.10.2-amzn-0	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.9.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.9.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Clasificaciones	Descripción	Acciones de reconfiguración
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
delta-defaults	Cambie los valores del archivo delta-defaults.conf de Delta.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j2	Cambiar la configuración de log4j2.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-lakeformation	Cambiar los valores en el archivo lakeformation.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-delta	Cambie los valores del archivo delta.properties de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Cambie los valores del archivo exchange-manager.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.9.0

Registro de cambios de la versión 6.9.0 y notas de la versión

Date	Evento	Descripción
30-08-2020	Actualización de las notas de la versión	Se ha agregado una corrección para el problema de desajuste de la secuencia de temporización
2023-08-21	Actualización de las notas de la versión	Se ha agregado un problema conocido con Hadoop 3.3.3.
2023-07-26	Actualización	Nuevas etiquetas de lanzamiento del sistema operativo `2.0.20230612.0` y `2.0.20230628.0` .
2022-12-13	Notas de la versión actualizadas	Se agregó una función y un problema conocido relacionados con el tiempo de ejecución con la IA SageMaker
29-11-2022	Se han actualizado las notas de la versión y la documentación	Se ha agregado la característica de integración de Amazon Redshift para Apache Spark
2022-11-23	Notas de la versión actualizadas	Se ha eliminado la entrada de Log4j
2022-11-18	Implementación completa	Amazon EMR 6.9 se ha implementado completamente en todas las regiones compatibles

Date	Evento	Descripción
2022-11-18	Publicación de documentos	Las notas de la versión 6.9 de Amazon EMR se publicaron por primera vez
2022-11-14	Versión inicial	Amazon EMR 6.9 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 6.8.1

Versiones de las aplicaciones de la versión 6.8.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, Phoenix, Pig, PrestoSpark Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.170

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.15.1	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.12	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.2
Hudi	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,273	0,273	0,272	0,267
Spark	3.3.0	3.3.0	3.2.1	3.2.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.9.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	388	378	367
Zeppelin	0.10.1	0.10.1	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.7	3.5.7

Notas de la versión 6.8.1

Las siguientes notas de la versión incluyen información sobre la versión 6.8.1 de Amazon EMR. Los cambios son respecto a la versión 6.8.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.8.1.

Cambios, mejoras y problemas resueltos

- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. Este enfoque también puede provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` se establece como `false` en `yarn-site.xml` para resolver este problema.

Si bien la corrección soluciona los problemas que ha introducido YARN-9608, es posible que los trabajos de Hive presenten errores debido a la pérdida de datos de mezclas aleatorias en los clústeres que tienen habilitado el escalado administrado. Se ha mitigado ese riesgo en esta versión al configurar también `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para las cargas de trabajo de Hive. Esta configuración solo está disponible con la versión 6.11.0 y posteriores de Amazon EMR.

- El recopilador de métricas no enviará ninguna métrica al plano de control después de una conmutación por error del nodo principal en los clústeres con la configuración de grupos de instancias.
- Esta versión elimina los reintentos de las solicitudes HTTP con errores a los puntos de conexión del recopilador de métricas.
- Esta versión incluye un cambio que permite que los clústeres de alta disponibilidad se recuperen de un estado de error tras el reinicio.
- Esta versión corrige un problema por el que las grandes excepciones creadas por los usuarios UIDs provocaban excepciones de desbordamiento.
- Esta versión corrige los problemas de tiempo de espera con el proceso de reconfiguración de Amazon EMR.
- Esta versión evita un problema en el que una reconfiguración con errores podría interrumpir otros procesos no relacionados.
- Esta versión incluye correcciones de seguridad.
- Esta versión corrige un problema por el que los clústeres que ejecutan cargas de trabajo en Spark con Amazon EMR podrían recibir silenciosamente resultados incorrectos con `contains`, `startsWith`, `endsWith` y `like`. Este problema se produce cuando utiliza las expresiones en campos particionados que tienen metadatos en Hive3 Metastore Server (HMS) de Amazon EMR.
- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.
- Es posible que Hive pierda datos si utiliza HDFS como directorio temporal y ha habilitado la combinación de archivos pequeños, y la tabla contiene rutas de partición estáticas.

- Esta versión corrige un problema de rendimiento con Hive si la combinación de archivos pequeños (deshabilitada de forma predeterminada) se habilita al final del trabajo de ETL.
- Esta versión corrige un problema con la limitación en Glue cuando no hay funciones definidas por el usuario (UDF).
- Esta versión corrige un problema que elimina los registros de contenedores mediante el servicio de agregación de registros de nodos antes de que el insertador de registros pudiera enviarlos a S3 en caso de retirada de YARN.
- Esta versión corrige la gestión de los archivos compactados o archivados con un seguimiento persistente de los archivos de almacenamiento. HBase
- Esta versión corrige un problema que afectaba al rendimiento de Spark cuando se establecía un valor `true` predeterminado para la configuración `spark.yarn.heterogeneousExecutors.enabled` en `spark-defaults.conf`.
- Esta versión corrige un problema que provocaba que la tarea de reducción no pudiera leer los datos de mezclas aleatorias. El problema provocaba errores en las consultas de Hive debido a un error de memoria dañada.
- Esta versión corrige un problema que provocaba un error en el aprovisionador de nodos si el servicio HDFS NameNode (NN) se bloqueaba en modo seguro durante la sustitución del nodo.
- Esta versión agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se detenga indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.
- Esta versión mejora la lógica de reducción vertical del clúster para que el clúster no intente reducir verticalmente los nodos principales por debajo de la configuración del factor de replicación de HDFS del clúster. Esto se ajusta a sus requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- El daemon de administración de registros se ha actualizado para identificar todos los registros que están en uso activo con controladores de archivos abiertos en el almacenamiento de instancias local y los procesos asociados. Esta actualización garantiza que Amazon EMR elimine correctamente los archivos y recupere espacio de almacenamiento una vez archivados los registros en Amazon S3.
- Esta versión incluye una mejora del daemon de administración de registros que elimina los directorios escalonados vacíos y no utilizados del sistema de archivos del clúster local. Un número

excesivamente elevado de directorios vacíos puede reducir el rendimiento de los daemons de Amazon EMR y provocar una sobreutilización del disco.

- Esta versión corrige un problema que podía producirse al crear un nodo de periferia cuando se replicaba uno de los nodos principales de un clúster con varios nodos principales. El nodo de periferia replicado podría provocar retrasos en las operaciones de reducción vertical o provocar un uso elevado de la memoria en los nodos principales. Para obtener más información sobre cómo crear un nodo perimetral para comunicarse con su clúster de EMR, consulte Edge Node Creator en el `aws-samples` repositorio de. GitHub
- Esta versión mejora el proceso de automatización que Amazon EMR utiliza para volver a montar los volúmenes de Amazon EBS en una instancia tras un reinicio.
- Esta versión corrige un problema que provocaba brechas intermitentes en las métricas de Hadoop que Amazon EMR publica en Amazon. CloudWatch
- Esta versión corrige un problema con los clústeres de EMR por el que se interrumpe una actualización del archivo de configuración de YARN que contiene la lista de nodos excluidos del clúster debido a una sobreutilización del disco. La actualización incompleta dificulta las futuras operaciones de reducción vertical de los clústeres. Esta versión garantiza que el clúster se mantenga en buen estado y que las operaciones de escalado funcionen según lo previsto.
- Esta versión mejora el daemon de administración de registros en el clúster para supervisar las carpetas de registro adicionales del clúster de EMR. Esta mejora minimiza los escenarios de sobreutilización del disco.
- Esta versión reinicia automáticamente el daemon de administración de registros en el clúster cuando se detiene. Esta mejora reduce el riesgo de que los nodos parezcan estar en mal estado debido a la sobreutilización del disco.
- Esta versión agrega compatibilidad con el archivado de registros en Amazon S3 durante la reducción vertical del clúster. Anteriormente, solo podía archivar archivos de registro en Amazon S3 durante la terminación del clúster. Esta nueva capacidad garantiza que los archivos de registro generados en el clúster persistan en Amazon S3 incluso después de terminar el nodo. Para más información, consulte Configurar el registro y la depuración de un clúster.
- Esta versión corrige un problema que se producía cuando el URI de Amazon S3 de una acción de arranque terminaba con un número de puerto, por ejemplo: `a.b.c.d:4345`. Amazon EMR los analizaba de forma incorrecta URIs, por lo que cualquier acción de arranque asociada fallaba.
- Esta versión corrige un problema de desajuste de la secuencia de tiempo entre Apache BigTop y Amazon EMR EC2 en la secuencia de inicio del clúster. Este desajuste de la secuencia de tiempo se produce cuando un sistema intenta llevar a cabo dos o más operaciones al mismo tiempo en lugar de hacerlas en la secuencia correcta. Como resultado, algunas configuraciones de clústeres

experimentaron tiempos de espera de inicio de instancias y tiempos de inicio de clústeres más lentos.

- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), AWS GovCloud (EEUU-Este), AWS

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud (EEUU-Oes te), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,14,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), China (Pekín), China (Ningxia), Oriente Medio (Emiratos Árabes Unidos), Asia Pacífico (Melbourne), Israel (Tel Aviv), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 001,0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14,328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Canadá (Central), Israel (Tel Aviv), (EEUU-Oeste) AWS GovCloud , (EEUU-Este) AWS GovCloud , China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

Versiones de los componentes de la versión 6.8.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.22.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR

Componente	Versión	Descripción
emrfs	2.53,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.15.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.15.1	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-8.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	3.2.1-amzn-8.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-8.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-8.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-8.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-8.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-8.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-8.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-8.1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	3.2.1-amzn-8.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-8.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.12-amzn-0.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.12-amzn-0.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.12-amzn-0.1	HBase cliente de línea de comandos.
hbase-rest-server	2.4.12-amzn-0.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.12-amzn-0.1	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.12-amzn-0.1	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-1.1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	3.1.3-amzn-1.1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-1.1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-1.1	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-1.1	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-1.1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-1.1	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.11.1-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.11.1-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.11.1-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.

Componente	Versión	Descripción
hudi-spark	0.11.1-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	0.14.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.7.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.2.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.2	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.273.3-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.273.3-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.273.3-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	388-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	388-amzn-0	Service para ejecutar partes de una consulta.
trino-client	388-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.0-amzn-0.1	Clientes de línea de comando de Spark.
spark-history-server	3.3.0-amzn-0.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.0-amzn-0.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.3.0-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.06.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.9.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.8.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.8.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Clasificaciones	Descripción	Acciones de reconfiguración
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Clasificaciones	Descripción	Acciones de reconfiguración
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.8.1

Registro de cambios de la versión 6.8.1 y notas de la versión

Date	Evento	Descripción
30-08-2020	Actualización de las notas de la versión	Se han agregado varias correcciones relacionadas con el plano de control a las notas de la versión
2023-08-21	Publicación de documentos	Las notas de la versión 6.8.1 de Amazon EMR se publicaron por primera vez
2023-08-16	Implementación completa	Amazon EMR 6.8.1 se ha implementado completamente en todas las regiones compatibles
2023-08-04	Versión inicial	Amazon EMR 6.8.1 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 6.8.0

Versiones de las aplicaciones de la versión 6.8.0

Esta versión incluye las siguientes aplicaciones: Flink,,,, Ganglia, HBase,, HCatalog, Hadoop, Hive,, Hudi, Hue, Iceberg,, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, OoziePhoenixPigPrestoSpark Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
AWS SDK for Java	1.12.170	1.12.170	1.12.170	1.12.31
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.14.2	1.14.2	1.14.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.2	3.1.2
Hudi	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1	0.12.0

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,273	0,272	0,267	0,261
Spark	3.3.0	3.2.1	3.2.0	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	378	367	360
Zeppelin	0.10.1	0.10.0	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.7	3.5.7	3.5.7

Notas de la versión 6.8.0

Las siguientes notas de la versión incluyen información sobre la versión 6.8.0 de Amazon EMR. Los cambios son respecto a la versión 6.7.0.

Nuevas características

- La función de pasos de Amazon EMR ahora es compatible con los clientes y puntos de conexión de Apache Livy. JDBC/ODBC Para más información, consulte Configuración de roles en tiempo de ejecución para los pasos de Amazon EMR.
- La versión 6.8.0 de Amazon EMR viene con la versión 2.4.12 de Apache HBase . Con esta HBase versión, puede archivar y eliminar sus tablas. HBase El proceso de archivado de Amazon S3 cambia el nombre de todos los archivos de tabla al directorio de archivado. Esto puede ser un proceso largo y costoso. Ahora puede omitir el proceso de archivado y rechazar y eliminar rápidamente tablas grandes. Para obtener más información, consulte Uso del HBase shell.

Problemas conocidos

- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. En Amazon EMR 6.8.0 y 6.9.0, este enfoque también podría provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  Con Amazon EMR 6.10.0, hay una solución alternativa para este problema: puede establecer el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` en `yarn-site.xml`. En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, la configuración se establece en `false` de manera predeterminada para resolver este problema.

Cambios, mejoras y problemas resueltos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leían las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, Amazon EMR producía un `NoSuchMethodError`. La versión 6.8.0 de Amazon EMR corrige este problema.
- Amazon EMR 6.8.0 viene con Apache Hudi 0.11.1; sin embargo, los clústeres de Amazon EMR 6.8.0 también son compatibles con el código abierto `hudi-spark3.3-bundle_2.12` de Hudi 0.12.0.
- La versión 6.8.0 de Amazon EMR incluye la versión 3.3.0 de Apache Spark. Esta versión de Spark utiliza Apache Log4j 2 y el archivo `log4j2.properties` para configurar Log4j en los procesos

de Spark. Si utiliza Spark en el clúster o crea clústeres de EMR con parámetros de configuración personalizados y desea actualizar a la versión 6.8.0 de Amazon EMR, debe migrar al nuevo formato de clave y clasificación de configuración `spark-log4j2` de Apache Log4j 2. Para obtener más información, consulte Migración de Apache Log4j 1.x a Log4j 2.x.

- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), AWS GovCloud (EEUU-Este), AWS

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud (EEUU-Oes te), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Europa (Zúrich)
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv)
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 119,1	4,14.301	3 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	22 de diciembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 912,1	4,14,291	6 de septiembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

Problemas conocidos

- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. Esto se debe a que Spark 3.2.0 establece

`spark.hadoopRDD.ignoreEmptySplits` en `true` de manera predeterminada. Como solución alternativa, establezca `spark.hadoopRDD.ignoreEmptySplits` explícitamente en `false`. La versión 6.9.0 de Amazon EMR corrige este problema.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Para obtener más información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Versiones de los componentes de la versión 6.8.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.2	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.7.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.22.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR
emrfs	2.53,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.15.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.15.1	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-8	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-8	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-8	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-8	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-8	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-8	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-8	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	3.2.1-amzn-8	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-8	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-8	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-8	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.12-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.12-amzn-0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.12-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.4.12-amzn-0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.12-amzn-0	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hbase-operator-tools	2.4.12-amzn-0	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-1	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.3-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hudi	0.11.1-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.11.1-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.11.1-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.11.1-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	0.14.0-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.9.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.

Componente	Versión	Descripción
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.7.0	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.2	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.273.3-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.273.3-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.273.3-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	388-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	388-amzn-0	Service para ejecutar partes de una consulta.
trino-client	388-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.3.0-amzn-0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	3.3.0-amzn-0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.3.0-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.3.0-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.06.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.9.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.5.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.8.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.8.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j2	Cambiar los valores en el archivo log4j2.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 6.8.0

Registro de cambios de la versión 6.8.0 y notas de la versión

Date	Evento	Descripción
2023-08-21	Actualización	Se ha agregado un problema conocido con Hadoop 3.3.3.
2023-07-26	Actualización	Nuevas etiquetas de lanzamiento del sistema operativo `2.0.20230612.0` y `2.0.20230628.0` .
2022-09-06	Implementación completa	Amazon EMR 6.8 se ha implementado completam ente en todas las regiones compatibles
2022-09-06	Publicación inicial	Las notas de la versión 6.8 de Amazon EMR se publicaron por primera vez
2022-08-31	Versión inicial	Amazon EMR 6.8 se ha implementado por primera vez en regiones comerciales limitadas

Amazon EMR, versión 6.7.0

Versiones de las aplicaciones de la versión 6.7.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
AWS SDK for Java	1.12.170	1.12.170	1.12.31	1.12.31
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.0	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.2	3.1.2	3.1.2
Hudi	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0
Hue	4.10.0	4.10.0	4.9.0	4.9.0

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
Iceberg	0.13.1-amzn-0	0.13.1	0.12.0	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,272	0,267	0,261	0,254.1
Spark	3.2.1	3.2.0	3.1.2	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	378	367	360	359
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.5.7

Notas de la versión 6.7.0

Las siguientes notas de la versión incluyen información sobre la versión 6.7.0 de Amazon EMR. Los cambios son respecto a la versión 6.6.0.

Fecha de lanzamiento inicial: 15 de julio de 2022

Nuevas características

- Amazon EMR ahora es compatible con Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272 y Trino 0.378.
- Admite controles de acceso basados en IAM Role y Lake Formation con pasos de EMR (Spark, Hive) para Amazon EMR en clústeres. EC2
- Admite las instrucciones de definición de datos de Apache Spark en clústeres habilitados para Apache Ranger. Esto ahora incluye compatibilidad para que las aplicaciones de Trino lean y escriban metadatos de Apache Hive en clústeres habilitados para Apache Ranger. Para más información, consulte Habilitar la gobernanza federada mediante Trino y Apache Ranger en Amazon EMR.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Oregón), EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney) AWS GovCloud), Israel (Tel Aviv), Canadá (Centro), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), (EE. UU. AWS GovCloud Este), AWS GovCloud (EEUU-Oes

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			te), China (Pekín), China (Ningxia), Israel (Tel Aviv)
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,14,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Europa (Zúrich)
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos Árabes

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Unidos), Europa (España), Europa (Zúrich)
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 131,0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 124,0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 218,0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 206,0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 116,0	4,14,328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 101,0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 020,1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 012,1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 906,0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 119,1	4,14.301	3 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	22 de diciembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 912,1	4,14,291	7 de octubre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2022 719,0	4,14,287	10 de agosto de 2022	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 606.1	4,14,281	15 de julio de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

Problemas conocidos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leen las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, se produce un error `NoSuchMethodError`

porque Amazon EMR utiliza un valor de `Hbase.compat.version` incorrecto. La versión 6.8.0 de Amazon EMR corrige este problema.

- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. Esto se debe a que Spark 3.2.0 establece `spark.hadoopRDD.ignoreEmptySplits` en `true` de manera predeterminada. Como solución alternativa, establezca `spark.hadoopRDD.ignoreEmptySplits` explícitamente en `false`. La versión 6.9.0 de Amazon EMR corrige este problema.
- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- La API `GetClusterSessionCredentials` no es compatible con los clústeres que se ejecutan en Amazon EMR 6.7 o versiones anteriores.

Versiones de los componentes de la versión 6.7.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.6.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.22.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR
emrfs	2.52,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.14.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.14.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-7	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-7	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-7	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-7	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-7	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-7	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	3.2.1-amzn-7	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-7	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-7	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-7	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-7	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.4-amzn-3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.4-amzn-3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.4-amzn-3	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	2.4.4-amzn-3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.4-amzn-3	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.4-amzn-3	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.3-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.3-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.3-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.3-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	3.1.3-amzn-0	Hive-hbase client.
hive-metastore-server	3.1.3-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server2	3.1.3-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.11.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.11.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.11.0-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.11.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	0.13.1-amzn-0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso

Componente	Versión	Descripción
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.0.194	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.2	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.272-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.272-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.272-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	378-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	378-amzn-0	Service para ejecutar partes de una consulta.
trino-client	378-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.2.1-amzn-0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	3.2.1-amzn-0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.2.1-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.2.1-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.02.0-amzn-1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.5.7	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.7.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.7.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 6.6.0

Versiones de las aplicaciones de la versión 6.6.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNet, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
AWS SDK for Java	1.12.170	1.12.31	1.12.31	1.11.977
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.0	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
HBase	2.4.4	2.4.4	2.4.4	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	0.13.1	0.12.0	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,261	0,254.1	0,245.1
Spark	3.2.0	3.1.2	3.1.2	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
Trino (PrestoSQL)	367	360	359	350
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.4.14

Notas de la versión 6.6.0

Las siguientes notas de la versión incluyen información sobre la versión 6.6.0 de Amazon EMR. Los cambios son respecto a la versión 6.5.0.

Fecha de lanzamiento inicial: 9 de mayo de 2022

Última actualización de la documentación: 15 de junio de 2022

Nuevas características

- Amazon EMR 6.6 ahora admite Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367 y PrestoDB 0.267.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), EE.UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este) AWS GovCloud , (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (Oregón), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia-Pacífico (Sídney), Canadá (Centro), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 414,0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 001,0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 223,0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 131,0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 124,0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 218,0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 206,0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 116,0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 101,0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 020,1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 012,1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 906,0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 119,1	4,14.301	3 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	22 de diciembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 912,1	4,14,291	7 de octubre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2022 805.0	4,14,287	30 de agosto de 2022	`us-west-1`

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 719,0	4,14,287	10 de agosto de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 426,0	4,14,281	10 de junio de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 406.1	4,14,275	2 de mayo de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

- Con Amazon EMR 6.6 y versiones posteriores, las aplicaciones que utilizan Log4j 1.x y Log4j 2.x se actualizan para utilizar Log4j 1.2.17 (o posterior) y Log4j 2.17.1 (o posterior), respectivamente, y no requieren el uso de las acciones de arranque proporcionadas para mitigar los problemas de CVE.
- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y

posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.

- A partir de Amazon EMR 5.32.0 y 6.5.0, el tamaño dinámico del ejecutor para Apache Spark se habilita de forma predeterminada. Para activar o desactivar esta característica, puede utilizar el parámetro de configuración `spark.yarn.heterogeneousExecutors.enabled`.

Cambios, mejoras y problemas resueltos

- Amazon EMR reduce el tiempo de inicio del clúster hasta 80 segundos de promedio para los clústeres que utilizan la opción de AMI predeterminada de EMR y que solo instalan aplicaciones comunes, como Apache Hadoop, Apache Spark y Apache Hive.

Problemas conocidos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leen las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, se produce un error `NoSuchMethodError` porque Amazon EMR utiliza un valor de `Hbase.compat.version` incorrecto. La versión 6.8.0 de Amazon EMR corrige este problema.
- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. Esto se debe a que Spark 3.2.0 establece `spark.hadoopRDD.ignoreEmptySplits` en `true` de manera predeterminada. Como solución alternativa, establezca `spark.hadoopRDD.ignoreEmptySplits` explícitamente en `false`. La versión 6.9.0 de Amazon EMR corrige este problema.
- En los clústeres de ejecución prolongada de Trino, Amazon EMR 6.6.0 habilita los parámetros de registro de recopilación de elementos no utilizados en el archivo jvm.config de Trino para obtener mejor información de los registros de recopilación de elementos no utilizados. Este cambio añade muchos registros de recolección de basura al archivo launcher.log (/var/log/trino/launcher.log). Si ejecuta clústeres de Trino en Amazon EMR 6.6.0, es posible que los nodos se queden sin espacio en disco después de que el clúster haya estado ejecutándose durante un par de días debido a los registros agregados.

La solución alternativa para este problema consiste en ejecutar el siguiente script como una acción de arranque para deshabilitar los parámetros de registro de recopilación de elementos no utilizados en jvm.config al crear o clonar el clúster para Amazon EMR 6.6.0.

```
#!/bin/bash
  set -ex
  PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
  sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versiones de los componentes de la versión 6.6.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.5.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.20.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR
emrfs	2.50,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.14.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.14.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-6	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-6	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	3.2.1-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.4-amzn-2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.4-amzn-2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.4-amzn-2	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	2.4.4-amzn-2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.4-amzn-2	Servicio que proporciona un punto final de Thrift para HBase.
hbase-operator-tools	2.4.4-amzn-2	Herramienta de reparació n para HBase clústeres de Apache.
hcatalog-client	3.1.2-amzn-7	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-7	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-7	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-7	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-7	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-7	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server2	3.1.2-amzn-7	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.10.1-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.10.1-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.10.1-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.10.1-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	0.13.1	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso

Componente	Versión	Descripción
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	11.0.194	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-connectors	5.1.2	Conectores de Apache Phoenix para Spark-3
phoenix-query-server	5.1.2	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.267-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.267-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.267-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	367-amzn-0	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	367-amzn-0	Service para ejecutar partes de una consulta.
trino-client	367-amzn-0	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.2.0-amzn-0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	3.2.0-amzn-0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.2.0-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.2.0-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	22.02.0-amzn-0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.5.7	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.6.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.6.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Cambiar los valores en el archivo iceberg.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 6.5.0

Versiones de las aplicaciones de la versión 6.5.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub,, Livy, MXNet, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
AWS SDK for Java	1.12.31	1.12.31	1.11.977	1.11.977
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.0	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
HBase	2.4.4	2.4.4	2.2.6	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.9.0
Iceberg	0.12.0	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,261	0,254.1	0,245.1	0,245.1
Spark	3.1.2	3.1.2	3.1.1	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
Trino (PrestoSQL)	360	359	350	350
Zeppelin	0.10.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.4.14	3.4.14

Notas de la versión 6.5.0

Las siguientes notas de la versión incluyen información sobre la versión 6.5.0 de Amazon EMR. Los cambios son respecto a la versión 6.4.0.

Fecha de lanzamiento inicial: 20 de enero de 2022

Fecha de lanzamiento: 21 de marzo de 2022

Nuevas características

- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.
- A partir de Amazon EMR 5.32.0 y 6.5.0, el tamaño dinámico del ejecutor para Apache Spark se habilita de forma predeterminada. Para activar o desactivar esta característica, puede utilizar el parámetro de configuración `spark.yarn.heterogeneousExecutors.enabled`.
- Compatibilidad con el formato de tabla abierto Apache Iceberg para conjuntos de datos analíticos muy grandes.
- Support para ranger-trino-plugin 2.0.1-amzn-1
- Compatibilidad con toree 0.5.0

Cambios, mejoras y problemas resueltos

- La versión 6.5 de Amazon EMR ahora es compatible con Apache Iceberg 0.12.0 y ofrece mejoras en el tiempo de ejecución con Tiempo de ejecución de Amazon EMR para Apache Spark, Tiempo de ejecución de Amazon EMR para Presto y Tiempo de ejecución de Amazon EMR para Apache Hive.
- Apache Iceberg es un formato de tabla abierto para grandes conjuntos de datos en Amazon S3 y proporciona un rendimiento de consultas rápido en tablas grandes, confirmaciones atómicas, escrituras simultáneas y evolución de tablas compatible con SQL. A partir de Amazon EMR 6.5, puede usar Apache Spark 3.1.2 con el formato de tabla de Iceberg.
- Apache Hudi 0.9 agrega compatibilidad con DDL y DML de Spark SQL. Esto le permite crear y upsert tablas de Hudi utilizando únicamente instrucciones SQL. Apache Hudi 0.9 también incluye mejoras en el rendimiento de las consultas y del escritor.
- El Tiempo de ejecución de Amazon EMR para Apache Hive mejora el rendimiento de Apache Hive en Amazon S3 al eliminar las operaciones de cambio de nombre durante las operaciones de preparación y mejora el rendimiento de los comandos de comprobación de metaalmacén (MSCK) que se utilizan para reparar tablas.

Problemas conocidos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leen las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, se produce un error `NoSuchMethodError` porque Amazon EMR utiliza un valor de `Hbase.compat.version` incorrecto. La versión 6.8.0 de Amazon EMR corrige este problema.
- Los clústeres de paquetes de HBase en alta disponibilidad (HA) no se aprovisionan con el tamaño de volumen y el tipo de instancia predeterminados. La solución alternativa para este problema consiste en aumentar el tamaño del volumen raíz.
- Para utilizar las acciones de Spark con Apache Oozie, debe agregar la siguiente configuración al archivo `workflow.xml` de Oozie. De lo contrario, varias bibliotecas críticas, como Hadoop y EMRFS, no aparecerán en la ruta de clases de los ejecutores de Spark que lance Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es

posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:

- Se analizan dos o más particiones de la misma tabla.
- Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
- El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 6.5.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.4.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.19.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR
emrfs	2.48,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.14.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.14.0	Administración de recursos en nodos EMR para Apache Flink. JobManager

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	3.2.1-amzn-5	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-5	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.4-amzn-1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	2.4.4-amzn-1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.4-amzn-1	HBase cliente de línea de comandos.
hbase-rest-server	2.4.4-amzn-1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.4-amzn-1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-6	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-6	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-6	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-6	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-6	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	3.1.2-amzn-6	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-6	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.9.0-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.9.0-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.9.0-amzn-1	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.9.0-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.9.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
iceberg	0.12.0	Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter

Componente	Versión	Descripción
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.1.2	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.261-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.261-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.261-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
trino-coordinator	360	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	360	Service para ejecutar partes de una consulta.
trino-client	360	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico

Componente	Versión	Descripción
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.1.2-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	3.1.2-amzn-1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.1.2-amzn-1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.1.2-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	0.4.1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.

Componente	Versión	Descripción
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.7	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.5.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.5.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
iceberg-defaults	Cambie los valores del archivo iceberg-defaults.conf de Iceberg.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Clasificaciones	Descripción	Acciones de reconfiguración
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 6.4.0

Versiones de las aplicaciones de la versión 6.4.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Trino, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
AWS SDK for Java	1.12.31	1.11.977	1.11.977	1.11.880
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.2.6	2.2.6	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	5.1.2	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,254.1	0,245.1	0,245.1	0,238,3
Spark	3.1.2	3.1.1	3.1.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	359	350	350	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.4.14	3.4.14	3.4.14

Notas de la versión 6.4.0

Las siguientes notas de la versión incluyen información sobre la versión 6.4.0 de Amazon EMR. Los cambios son respecto a la versión 6.3.0.

Fecha de lanzamiento inicial: 20 de septiembre de 2021

Fecha de lanzamiento: 21 de marzo de 2022

Aplicaciones compatibles

- AWS SDK para Java versión 1.12.31
- CloudWatch Sink versión 2.2.0
- Conector de DynamoDB, versión 4.16.0
- EMRFS, versión 2.47.0
- Amazon EMR Goodies, versión 3.2.0
- Conector de Kinesis para Amazon EMR, versión 3.5.0
- Servidor de registros de Amazon EMR, versión 2.1.0
- Scripts de Amazon EMR, versión 2.5.0
- Flink, versión 1.13.1
- Versión de Ganglia 3.7.2
- AWS Cliente Glue Hive Metastore versión 3.3.0
- Hadoop, versión 3.2.1-amzn-4
- HBase versión 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog versión 3.1.2-amzn-5
- Hive, versión 3.1.2-amzn-5
- Hudi versión 0.8.0-amzn-0
- Hue versión 4.9.0
- Java JDK, versión Corretto-8.302.08.1 (compilación 1.8.0_302-b08)
- JupyterHub versión 1.4.1
- Livy versión 0.7.1-incubating
- MXNet versión 1.8.0
- Oozie versión 5.2.1
- Phoenix versión 5.1.2
- Pig, versión 0.17.0
- Presto, versión 0.254.1-amzn-0
- Trino, versión 359

- Apache Ranger KMS (cifrado transparente multimaestro), versión 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker Spark SDK versión 1.4.1
- Scala, versión 2.12.10 (máquina virtual de servidor OpenJDK de 64 bits, Java 1.8.0_282)
- Spark, versión 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop, versión 1.4.7
- TensorFlow versión 2.4.1
- Tez, versión 0.9.2
- Zeppelin versión 0.9.0
- ZooKeeper versión 3.5.7
- Conectores y controladores: conector de DynamoDB 4.16.0

Nuevas características

- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.
- En los clústeres de Amazon EMR habilitados para Apache Ranger, puede utilizar Apache Spark SQL para insertar datos o actualizar las tablas del metaalmacén de Apache Hive mediante `INSERT INTO`, `INSERT OVERWRITE` y `ALTER TABLE`. Al utilizar ALTER TABLE con Spark SQL, la ubicación de una partición debe ser el directorio secundario de la ubicación de una tabla. Amazon EMR no admite actualmente la inserción de datos en una partición en la que la ubicación de la partición es diferente de la ubicación de la tabla.
- PrestoSQL se ha renombrado como Trino.
- Hive: la ejecución de consultas SELECT sencillas con la cláusula LIMIT se acelera al detener la ejecución de las consultas en cuanto se obtiene el número de registros mencionado en la cláusula LIMIT. Las consultas SELECT simples son consultas que no tienen GROUP BY u ORDER BY por

cláusula o consultas que no tienen una etapa reductora. Por ejemplo, `SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`.

Control de simultaneidad de Hudi

- Hudi ahora es compatible con el control de simultaneidad optimista (OCC), que se puede utilizar con operaciones de escritura como UPSERT e INSERT para permitir cambios desde varios escritores a la misma tabla de Hudi. Se trata de un OCC de nivel de archivo, por lo que dos confirmaciones (o escritores) pueden escribir en la misma tabla, siempre que sus cambios no entren en conflicto. Para más información, consulte Control de simultaneidad de Hudi.
- Los clústeres de Amazon EMR tienen instalado ZooKeeper, que se puede utilizar como proveedor de bloqueos para OCC. Para facilitar el uso de esta característica, los clústeres de Amazon EMR tienen las siguientes propiedades preconfiguradas:

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

Para habilitar el OCC, debe configurar las siguientes propiedades con sus opciones de trabajo de Hudi o en el nivel de clúster mediante la API de configuraciones de Amazon EMR:

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Hudi Monitoring: CloudWatch integración de Amazon para informar sobre las métricas de Hudi

- Amazon EMR admite la publicación de Hudi Metrics en Amazon. CloudWatch Se habilita mediante el establecimiento de las siguientes configuraciones obligatorias:

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- Las siguientes son configuraciones opcionales de Hudi que se pueden cambiar:

Opción	Descripción	Valor
hoodie.metrics.cloudwatch.report.period.seconds	Frecuencia (en segundos) con la que se reportan las métricas a Amazon CloudWatch	El valor predeterminado es 60 segundos, lo cual está bien para la resolución predeterminada de un minuto que ofrece Amazon. CloudWatch
hoodie.metrics.cloudwatch.metric.prefix	Prefijo que se agregará al nombre de cada métrica	El valor predeterminado es vacío (sin prefijo)
hoodie.metrics.cloudwatch.namespace	Espacio de CloudWatch nombres de Amazon en el que se publican las métricas	El valor predeterminado es Hudi
hoodie.metrics.cloudwatch.maxDatumsPerSolicitud	Número máximo de datos que se pueden incluir en una solicitud a Amazon CloudWatch	El valor predeterminado es 20, que es el mismo que el CloudWatch valor predeterminado de Amazon

Soporte y mejoras en las configuraciones de Amazon EMR Hudi

- Los clientes ahora pueden aprovechar la API de configuraciones de EMR y la característica de reconfiguración para configurar los ajustes de Hudi en el nivel de clúster. Se ha introducido un nuevo soporte de configuración basado en archivos a través de/etc/hudi/conf/hudi-defaults.conf, similar al de otras aplicaciones como Spark, Hive, etc. EMR configura algunos valores predeterminados para mejorar la experiencia del usuario:

  — `hoodie.datasource.hive_sync.jdbcurl` se ha configurado en la URL del servidor de Hive del clúster y ya no es necesario especificarlo. Esto resulta especialmente útil cuando se ejecuta un trabajo en el modo de clúster de Spark, en el que anteriormente había que especificar la IP maestra de Amazon EMR.

  — configuraciones HBase específicas, que son útiles para usar el índice con Hudi. HBase

— Configuración específica para el proveedor de bloqueos de ZooKeeper, tal como se explica en la sección de control de simultaneidad, que facilita el uso del control de simultaneidad optimista (OCC).

- Se han introducido cambios adicionales para reducir el número de configuraciones que hay que pasar y, en la medida de lo posible, deducir automáticamente:

  — La palabra clave `partitionBy` se puede utilizar para especificar la columna de particiones.

  — Al activar Hive Sync, ya no es obligatorio pasar `HIVE_TABLE_OPT_KEY`, `HIVE_PARTITION_FIELDS_OPT_KEY`, `HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`. Estos valores se pueden deducir del nombre de la tabla de Hudi y del campo de partición.

  — No es obligatorio pasar `KEYGENERATOR_CLASS_OPT_KEY` y se puede deducir de casos más simples de `SimpleKeyGenerator` y `ComplexKeyGenerator`.

Advertencias de Hudi

- Hudi no admite la ejecución vectorizada en las tablas Hive for Merge on Read (MoR) y Bootstrap. Por ejemplo, `count(*)` presenta un error con la tabla en tiempo real de Hudi cuando `hive.vectorized.execution.enabled` se establece en true. Como solución alternativa, puede deshabilitar la lectura vectorizada al establecer `hive.vectorized.execution.enabled` en `false`.
- La compatibilidad con varios escritores no es compatible con la característica de arranque de Hudi.
- Flink Streamer y Flink SQL son características experimentales de esta versión. Se recomienda el uso de estas características en implementaciones de producción.

Cambios, mejoras y problemas resueltos

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

- Anteriormente, el reinicio manual del administrador de recursos en un clúster multimaestro provocaba que los daemons en el clúster de Amazon EMR, como ZooKeeper, recargaran todos los nodos perdidos o retirados previamente en el archivo znode de ZooKeeper. Esto provocaba que se superaran los límites predeterminados en ciertas situaciones. Amazon EMR ahora elimina del archivo de ZooKeeper los registros de nodos perdidos o retirados que tengan más de una hora de antigüedad y se han incrementado los límites internos.

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Configuración de un clúster para corregir los problemas de rendimiento del servidor YARN Timeline de Apache en las versiones 1 y 1.5

  Las versiones 1 y 1.5 del servidor YARN Timeline de Apache pueden provocar problemas de rendimiento con clústeres de EMR grandes y muy activos, especialmente con `yarn.resourcemanager.system-metrics-publisher.enabled=true`, que es la configuración predeterminada de Amazon EMR. Un servidor YARN Timeline v2 de código abierto resuelve el problema de rendimiento relacionado con la escalabilidad del servidor YARN Timeline.

  Otras alternativas para este problema incluyen:

  - Configuración de yarn.resourcemanager. system-metrics-publisher.enabled=false en yarn-site.xml.
  - La habilitación de la corrección para este problema al crear un clúster, tal y como se describe a continuación.

Las siguientes versiones de Amazon EMR contienen una corrección para este problema de rendimiento del servidor YARN Timeline.

EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

Para habilitar la corrección en cualquiera de las versiones de Amazon EMR especificadas anteriormente, defina estas propiedades como `true` en un archivo JSON de configuración que se pasa mediante el parámetro de comando `aws emr create-cluster`: `--configurations file://./configurations.json`. También puede habilitar la corrección mediante la interfaz de usuario de la consola de reconfiguración.

Ejemplo del contenido del archivo configurations.json:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- Los servidores WebHDFS y HttpFS están deshabilitados de forma predeterminada. Puede volver a habilitar WebHDFS mediante la configuración de Hadoop, `dfs.webhdfs.enabled`. El servidor HttpFS se puede iniciar mediante `sudo systemctl start hadoop-httpfs`.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de discusión. AWS Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.
- Hive: se ha mejorado el rendimiento de las consultas de escritura al permitir el uso de un directorio temporal en HDFS para el último trabajo. Los datos temporales del trabajo final se escriben en HDFS en lugar de Amazon S3 y el rendimiento mejora porque los datos se mueven de HDFS a la ubicación de la tabla final (Amazon S3) en lugar de entre dispositivos Amazon S3.

- Hive: mejora el tiempo de compilación de consultas hasta 2,5 veces con la eliminación de particiones del metaalmacén de Glue.
- De forma predeterminada, cuando Hive UDFs transfiere los elementos integrados al servidor Hive Metastore, solo un subconjunto de los integrados UDFs se pasa a Glue Metastore, ya que Glue solo admite operadores de expresiones limitadas. Si configura `hive.glue.partition.pruning.client=true`, toda la eliminación de particiones se realiza en el cliente. Si configura `hive.glue.partition.pruning.server=true`, entonces toda la eliminación de particiones se realiza en el lado del servidor.

Problemas conocidos

- Las consultas de Hue no funcionan en Amazon EMR 6.4.0 porque el servidor HttpFS de Apache Hadoop está deshabilitado de forma predeterminada. Para usar Hue en Amazon EMR 6.4.0, inicie manualmente el servidor HttpFS en el nodo principal de Amazon EMR mediante `sudo systemctl start hadoop-httpfs` o utilice un paso de Amazon EMR.
- La característica Cuadernos de Amazon EMR que se utiliza con la suplantación de usuarios de Livy no funciona porque HttpFS está deshabilitado de forma predeterminada. En este caso, el cuaderno de EMR no puede conectarse al clúster que tiene habilitada la suplantación de Livy. La solución alternativa consiste en iniciar el servidor HttpFS antes de conectar el cuaderno de EMR al clúster mediante `sudo systemctl start hadoop-httpfs`.
- En la versión 6.4.0 de Amazon EMR, Phoenix no admite el componente de conectores de Phoenix.
- Para utilizar las acciones de Spark con Apache Oozie, debe agregar la siguiente configuración al archivo `workflow.xml` de Oozie. De lo contrario, varias bibliotecas críticas, como Hadoop y EMRFS, no aparecerán en la ruta de clases de los ejecutores de Spark que lance Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.

- El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter / (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 6.4.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.3.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.18.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR
emrfs	2.47,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.13.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.13.1	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-4	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	3.2.1-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.4.4-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.4.4-amzn-0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.4.4-amzn-0	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	2.4.4-amzn-0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.4.4-amzn-0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-5	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-5	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-5	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-5	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-5	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-5	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-5	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hudi	0.8.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.8.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-trino	0.8.0-amzn-0	Biblioteca de paquetes para ejecutar Trino con Hudi.
hudi-spark	0.8.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.9.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.1.2	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.1.2	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.254.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.254.1-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.254.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
trino-coordinator	359	Servicio para aceptar consultas y administrar la ejecución de consultas entre trino-workers.
trino-worker	359	Service para ejecutar partes de una consulta.
trino-client	359	Cliente de línea de comandos Trino que se instala en los maestros en espera de un clúster de HA donde el servidor Trino no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.1.2-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.1.2-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.1.2-amzn-0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.1.2-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	0.4.1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.5.7	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.5.7	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.4.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.4.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Clasificaciones	Descripción	Acciones de reconfiguración
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)

Clasificaciones	Descripción	Acciones de reconfiguración
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
trino-log	Cambiar los valores en el archivo log.properties de Trino.	Restarts Trino-Server (for Trino)

Clasificaciones	Descripción	Acciones de reconfiguración
trino-config	Cambiar los valores en el archivo config.properties de Trino.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Trino.	Restarts Trino-Server (for Trino)
trino-env	Cambie los valores del archivo trino-env.sh de Trino.	Restarts Trino-Server (for Trino)
trino-node	Cambiar los valores en el archivo node.properties de Trino.	Not available.
trino-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Trino.	Not available.
trino-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Trino.	Not available.
trino-connector-hive	Cambiar los valores en el archivo hive.properties de Trino.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Cambiar los valores en el archivo jmx.properties de Trino.	Not available.
trino-connector-kafka	Cambiar los valores en el archivo kafka.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-localfile	Cambiar los valores en el archivo localfile.properties de Trino.	Not available.
trino-connector-memory	Cambiar los valores en el archivo memory.properties de Trino.	Not available.
trino-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Trino.	Not available.
trino-connector-mysql	Cambiar los valores en el archivo mysql.properties de Trino.	Not available.
trino-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Trino.	Not available.
trino-connector-raptor	Cambiar los valores en el archivo raptor.properties de Trino.	Not available.
trino-connector-redis	Cambiar los valores en el archivo redis.properties de Trino.	Not available.
trino-connector-redshift	Cambiar los valores en el archivo redshift.properties de Trino.	Not available.
trino-connector-tpch	Cambiar los valores en el archivo tpch.properties de Trino.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
trino-connector-tpcds	Cambiar los valores en el archivo tpcds.properties de Trino.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 6.3.1

Versiones de las aplicaciones de la versión 6.3.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, Phoenix, Pig, PrestoPrestoSQLSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
AWS SDK for Java	1.11.977	1.11.977	1.11.880	1.11.880
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245.1	0,245.1	0,238,3	0,238,3
Spark	3.1.1	3.1.1	3.0.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	350	350	343	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.3.1

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

Cambios, mejoras y problemas resueltos

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.

- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de discusión. AWS Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a  b`.

- El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter / (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 6.3.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.2.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.18.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR
emrfs	2.46.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.12.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.12.1	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-3.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-3.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-3.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-3.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-3.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-3.1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	3.2.1-amzn-3.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-3.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-3.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-3.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-3.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.2.6-amzn-1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.2.6-amzn-1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.6-amzn-1	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	2.2.6-amzn-1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.6-amzn-1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-4	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-4	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hudi	0.7.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.7.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-prestosql	0.7.0-amzn-0	Biblioteca de paquetes para ejecutar PrestoSQL con Hudi.
hudi-spark	0.7.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.9.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.2.2	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.0.0- -2.0 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.245.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.245.1-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.245.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
prestosql-coordinator	350	Servicio para aceptar las consultas y administrar la ejecución de consultas entre prestosql-workers.
prestosql-worker	350	Service para ejecutar partes de una consulta.
prestosql-client	350	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.1.1-amzn-0.1	Clientes de línea de comando de Spark.
spark-history-server	3.1.1-amzn-0.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.1.1-amzn-0.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.1.1-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	0.4.1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.3.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.3.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop

Clasificaciones	Descripción	Acciones de reconfiguración
		Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
prestosql-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Cambiar los valores en el archivo node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Cambiar los valores en el archivo hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Cambiar los valores en el archivo jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Cambiar los valores en el archivo kafka.properties de PrestoSQL.	Not available.
prestosql-connector-localfile	Cambiar los valores en el archivo localfile.properties de PrestoSQL.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-connector-memory	Cambiar los valores del archivo memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Cambiar los valores en el archivo mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Cambiar los valores en el archivo raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Cambiar los valores en el archivo redis.properties de PrestoSQL.	Not available.
prestosql-connector-redshift	Cambiar los valores en el archivo redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Cambiar los valores en el archivo tpch.properties de PrestoSQL.	Not available.
prestosql-connector-tpcds	Cambiar los valores del archivo tpcds.properties de PrestoSQL.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServe r2.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 6.3.0

Versiones de las aplicaciones de la versión 6.3.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, Phoenix, Pig, PrestoPrestoSQLSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
AWS SDK for Java	1.11.977	1.11.880	1.11.880	1.11,828
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245.1	0,238,3	0,238,3	0.232
Spark	3.1.1	3.0.1	3.0.1	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	350	343	343	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.3.0

Las siguientes notas de la versión incluyen información sobre la versión 6.3.0 de Amazon EMR. Los cambios son respecto a la versión 6.2.0.

Fecha de lanzamiento inicial: 12 de mayo de 2021

Fecha de la última actualización: 9 de agosto de 2021

Aplicaciones compatibles

- AWS SDK para Java versión 1.11.977

- CloudWatch Sink versión 2.1.0
- Conector de DynamoDB, versión 4.16.0
- EMRFS, versión 2.46.0
- Amazon EMR Goodies, versión 3.2.0
- Conector de Kinesis para Amazon EMR, versión 3.5.0
- Servidor de registros de Amazon EMR, versión 2.0.0
- Scripts de Amazon EMR, versión 2.5.0
- Flink, versión 1.12.1
- Versión de Ganglia 3.7.2
- AWS Cliente Glue Hive Metastore versión 3.2.0
- Hadoop, versión 3.2.1-amzn-3
- HBase versión 2.2.6-amzn-1
- HBase-operator-tools 1.0.0
- HCatalog versión 3.1.2-amzn-0
- Hive, versión 3.1.2-amzn-4
- Hudi versión 0.7.0-amzn-0
- Hue versión 4.9.0
- Java JDK, versión Corretto-8.282.08.1 (compilación 1.8.0_282-b08)
- JupyterHub versión 1.2.0
- Livy versión 0.7.0-incubating
- MXNet versión 1.7.0
- Oozie versión 5.2.1
- Versión de Phoenix 5.0.0
- Pig, versión 0.17.0
- Presto, versión 0.245.1-amzn-0
- PrestoSQL, versión 350
- Apache Ranger KMS (cifrado transparente multimaestro), versión 2.0.0
- ranger-plugins 2.0.1-amzn-0

- ranger-s3-plugin 1.1.0
- SageMaker Spark SDK versión 1.4.1
- Scala, versión 2.12.10 (máquina virtual de servidor OpenJDK de 64 bits, Java 1.8.0_282)
- Spark, versión 3.1.1-amzn-0
- spark-rapids 0.4.1
- Sqoop, versión 1.4.7
- TensorFlow versión 2.4.1
- Tez, versión 0.9.2
- Zeppelin versión 0.9.0
- Versión de Zookeeper 3.4.14
- Conectores y controladores: conector de DynamoDB 4.16.0

Nuevas características

- Amazon EMR es compatible con Puntos de acceso de Amazon S3, una característica de Amazon S3 que le permite administrar fácilmente el acceso a los lagos de datos compartidos. Con su alias de Puntos de acceso de Amazon S3, puede simplificar el acceso a los datos a escala en Amazon EMR. Puede utilizar los puntos de acceso Amazon S3 con todas las versiones de Amazon EMR sin coste adicional en todas AWS las regiones en las que Amazon EMR esté disponible. Para más información acerca de los puntos de acceso de Amazon S3 y los alias de punto de acceso, consulte Uso de un alias de estilo bucket para el punto de acceso en la Guía del usuario de Amazon S3.
- Los parámetros nuevos `DescribeReleaseLabel` y `ListReleaseLabel` de la API proporcionan detalles de la etiqueta de la versión de Amazon EMR. Puede enumerar mediante programación las versiones disponibles en la región en la que se ejecuta la solicitud de API y enumerar las aplicaciones disponibles para una etiqueta de versión específica de Amazon EMR. Los parámetros de las etiquetas de versión también enumeran las versiones de Amazon EMR compatibles con una aplicación específica, como Spark. Esta información se puede utilizar para iniciar clústeres de Amazon EMR mediante programación. Por ejemplo, puede iniciar un clúster con la versión más reciente de los resultados de `ListReleaseLabel`. Para obtener más información, consulte DescribeReleaseLabely ListReleaseLabelsen la referencia de la API de Amazon EMR.
- A partir de Amazon EMR 6.3.0, puede iniciar un clúster que se integre de forma nativa con Apache Ranger. Apache Ranger es un marco de código abierto para habilitar, supervisar y administrar la seguridad integral de los datos en toda la plataforma Hadoop. Para obtener más información,

consulte Apache Ranger. Con la integración nativa, puede utilizar su propio Apache Ranger para aplicar un control de acceso a los datos detallado en Amazon EMR. Consulte Integrar Amazon EMR con Apache Ranger en la Guía de administración de Amazon EMR.

- Políticas gestionadas con ámbito de aplicación: para ajustarse a las prácticas AWS recomendadas, Amazon EMR ha introducido la versión 2 de las políticas gestionadas predeterminadas con ámbito de EMR como sustitutivas de las políticas que quedarán en desuso. Consulte Políticas administradas por Amazon EMR.
- Estado de soporte de Instance Metadata Service (IMDS) V2: para Amazon EMR 6.2 o posterior, los componentes de Amazon EMR se IMDSv2 utilizan para todas las llamadas de IMDS. Para las llamadas de IMDS en el código de su aplicación, puede utilizar ambos IMDSv1 o configurar el IMDS para que se utilice únicamente con fines de seguridad adicionales. IMDSv2 IMDSv2 Si la deshabilita IMDSv1 en versiones anteriores de Amazon EMR 6.x, se produce un error al iniciar el clúster.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.

- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- El modo de explicación de la interfaz de usuario de Spark SQL ha cambiado de `extended` a `formatted` en Spark 3.1. Amazon EMR lo revirtió a `extended` para incluir información del plan lógico en la interfaz de usuario de Spark SQL. Esto se puede revertir al establecer `spark.sql.ui.explainMode` en `formatted`.
- Se ha agregado portabilidad con versiones anteriores para las siguientes confirmaciones desde la ramificación maestra de Spark.

  - [SPARK-34752] [BUILD] Actualizar Jetty a la versión 9.4.37 para abordar CVE-2020-27223.

  - [SPARK-34534] Corrija el orden de BlockIDs cuando se utiliza para buscar bloques. FetchShuffleBlocks

  - [SPARK-34681] [SQL] Corregir un error de combinación hash desordenada externa completa al crear el lado izquierdo con condiciones desiguales.

  - [SPARK-34497] [SQL] Corregir los proveedores de conexión JDBC integrados para restaurar los cambios en el contexto de seguridad de JVM.
- Para mejorar la interoperabilidad con el RAPIDs plugin Nvidia Spark, se ha añadido una solución alternativa para solucionar un problema que impedía que se activara la reducción dinámica de particiones cuando se utilizaba Nvidia Spark RAPIDs con la ejecución de consultas adaptativas desactivada, consulte los números #1378 y #1386 de RAPIDS. Para más información sobre la nueva configuración `spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`, consulte Problema núm. 1386 de RAPIDS.
- En Spark 3.1 de código abierto, se ha cambiado el algoritmo predeterminado del confirmador de salida de archivos de la versión 2 a la versión 1. Para más información, consulte Amazon EMR optimiza el rendimiento de Spark: eliminación dinámica de particiones.

- Amazon EMR volvió a la versión 2 del algoritmo, la predeterminada que se utilizaba en las versiones anteriores de Amazon EMR 6.x, para evitar la regresión del rendimiento. Para restaurar el comportamiento de Spark 3.1 de código abierto, establezca `spark.hadoop.mapreduce.fileoutputcommitter.algorithm.version` en `1`. Spark, de código abierto, ha realizado este cambio porque la confirmación de tareas en la versión 2 del algoritmo de confirmación de salida de archivos no es atómica, lo que puede provocar un problema de corrección de los datos de salida en algunos casos. Sin embargo, la confirmación de tareas en la versión 1 del algoritmo tampoco es atómica. En algunos escenarios, la confirmación de tareas incluye una eliminación realizada antes del cambio de nombre. Esto puede provocar un problema silencioso de corrección de los datos.
- Se corrigieron los problemas de escalado administrado en versiones anteriores de Amazon EMR y se realizaron mejoras para reducir significativamente las tasas de errores de las aplicaciones.
- Se instaló el paquete de SDK de AWS Java en cada clúster nuevo. Se trata de un contenedor único que contiene todos los servicios SDKs y sus dependencias, en lugar de contenedores de componentes individuales. Para más información, consulte Java SDK Bundled Dependency.

Problemas conocidos

- En el caso de los clústeres de subredes privadas de Amazon EMR 6.3.0 y 6.2.0, no puede acceder a la interfaz de usuario web de Ganglia. Aparecerá el error “Acceso denegado (403)”. Otras páginas web UIs, como Spark, Hue, Zeppelin JupyterHub, Livy y Tez, funcionan con normalidad. El acceso a la interfaz de usuario web de Ganglia en los clústeres de subredes públicas también funciona con normalidad. Para resolver este problema, reinicie el servicio httpd en el nodo principal con `sudo systemctl restart httpd`. Este problema se ha corregido en la versión 6.4.0 de Amazon EMR.
- Cuando el catálogo de datos de AWS Glue está activado, es posible que no se pueda utilizar Spark para acceder a una base de datos de AWS Glue con un URI de ubicación de cadena nula. Esto ocurre con las versiones anteriores de Amazon EMR, excepto con el SPARK-31709 (https://issues.apache). org/jira/browse/SPARK-31709) hace que se aplique a más casos. Por ejemplo, al crear una tabla en la base de datos AWS Glue predeterminada cuyo URI de ubicación es una cadena nula, se produce un `spark.sql("CREATE TABLE mytest (key string) location '/table_path';")` error con el mensaje «No se puede crear una ruta a partir de una cadena vacía». Para solucionar este problema, establece manualmente el URI de ubicación de tus bases de datos de AWS Glue y, a continuación, crea tablas dentro de estas bases de datos con Spark.
- En Amazon EMR 6.3.0, PrestoSQL se ha actualizado de la versión 343 a la versión 350. Hay dos cambios respecto a la seguridad en el código abierto que se relacionan con este cambio de

versión. El control de acceso al catálogo basado en archivos cambia de `deny` a `allow` cuando las reglas de propiedades de la tabla, el esquema o la sesión no están definidas. Además, el control de acceso al sistema basado en archivos se ha modificado para admitir archivos sin reglas de catálogo definidas. En este caso, se permite todo el acceso a los catálogos.

Para más información, consulte Release 344 (9 Oct 2020).

- Ten en cuenta el directorio de usuarios de Hadoop (/) home/hadoop) is readable by everyone. It has Unix 755 (drwxr-xr-x) directory permissions to allow read access by frameworks like Hive. You can put files in /home/hadoop y sus subdirectorios, pero ten en cuenta los permisos de esos directorios para proteger la información confidencial.
- Reducir el límite de «máximo de archivos abiertos» en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

  Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

  Establecimiento explícito de un ulimit desde la línea de comandos

  1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

2. Reinicie InstanceController

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no

descargan ni instalan automáticamente esas actualizaciones. Para obtener actualizaciones del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.

- Para utilizar las acciones de Spark con Apache Oozie, debe agregar la siguiente configuración al archivo `workflow.xml` de Oozie. De lo contrario, varias bibliotecas críticas, como Hadoop y EMRFS, no aparecerán en la ruta de clases de los ejecutores de Spark que lance Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre `a` y `b` en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 6.3.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.2.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.18.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.1.0	Conector S3Select de EMR
emrfs	2.46.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
flink-client	1.12.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.12.1	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-3	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	3.2.1-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-3	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	2.2.6-amzn-1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.2.6-amzn-1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.6-amzn-1	HBase cliente de línea de comandos.
hbase-rest-server	2.2.6-amzn-1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.6-amzn-1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-4	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	3.1.2-amzn-4	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.7.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.7.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-prestosql	0.7.0-amzn-0	Biblioteca de paquetes para ejecutar PrestoSQL con Hudi.
hudi-spark	0.7.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.9.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.2.2	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de base de datos de MariaDB.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.0.0- -2.0 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.245.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.245.1-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.245.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
prestosql-coordinator	350	Servicio para aceptar las consultas y administrar la ejecución de consultas entre prestosql-workers.
prestosql-worker	350	Service para ejecutar partes de una consulta.
prestosql-client	350	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.1.1-amzn-0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	3.1.1-amzn-0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.1.1-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.1.1-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	0.4.1	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.3.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.3.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Restarts Flink history server.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
prestosql-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Cambiar los valores en el archivo node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Cambiar los valores en el archivo hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-connector-jmx	Cambiar los valores en el archivo jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Cambiar los valores en el archivo kafka.properties de PrestoSQL.	Not available.
prestosql-connector-localfile	Cambiar los valores en el archivo localfile.properties de PrestoSQL.	Not available.
prestosql-connector-memory	Cambiar los valores del archivo memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Cambiar los valores en el archivo mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Cambiar los valores en el archivo raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Cambiar los valores en el archivo redis.properties de PrestoSQL.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-connector-redshift	Cambiar los valores en el archivo redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Cambiar los valores en el archivo tpch.properties de PrestoSQL.	Not available.
prestosql-connector-tpcds	Cambiar los valores del archivo tpcds.properties de PrestoSQL.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie and HiveServer2.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.pro perties.	Restarts Zookeeper server.

Amazon EMR, versión 6.2.1

Versiones de las aplicaciones de la versión 6.2.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, Phoenix, Pig, PrestoPrestoSQLSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
AWS SDK for Java	1.11.880	1.11.880	1.11,828	1.11.828
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5	2.2.5
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
Hudi	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0,238,3	0.232	0.232
Spark	3.0.1	3.0.1	3.0.0	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	343	338	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.2.1

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

Cambios, mejoras y problemas resueltos

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket

de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de AWS debate. Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 6.2.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*-

`amzn-EmrVersion`. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.0.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.16.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.0.0	Conector S3Select de EMR
emrfs	2.44.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.

Componente	Versión	Descripción
flink-jobmanager-config	1.11.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-2.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-2.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-2.1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	3.2.1-amzn-2.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-2.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-2.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-2.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-2.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-2.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-2.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-2.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	2.2.6-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.2.6-amzn-0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.6-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.2.6-amzn-0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.6-amzn-0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-3	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-3	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-3	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-3	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	3.1.2-amzn-3	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-3	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-3	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.6.0-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.6.0-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-prestosql	0.6.0-amzn-1	Biblioteca de paquetes para ejecutar PrestoSQL con Hudi.
hudi-spark	0.6.0-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.8.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64+	Servidor de base de datos de MariaDB.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.0.0- -2.0 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.238.3-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.238.3-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.238.3-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
prestosql-coordinator	343	Servicio para aceptar las consultas y administrar la ejecución de consultas entre prestosql-workers.
prestosql-worker	343	Service para ejecutar partes de una consulta.
prestosql-client	343	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.0.1-amzn-0.1	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	3.0.1-amzn-0.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.0.1-amzn-0.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.0.1-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	0.2.0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.3.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0-preview1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.2.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.2.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Not available.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Not available.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Not available.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS ZKFC.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
prestosql-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Cambiar los valores en el archivo node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Cambiar los valores en el archivo hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-connector-jmx	Cambiar los valores en el archivo jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Cambiar los valores en el archivo kafka.properties de PrestoSQL.	Not available.
prestosql-connector-localfile	Cambiar los valores en el archivo localfile.properties de PrestoSQL.	Not available.
prestosql-connector-memory	Cambiar los valores del archivo memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Cambiar los valores en el archivo mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Cambiar los valores en el archivo raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Cambiar los valores en el archivo redis.properties de PrestoSQL.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-connector-redshift	Cambiar los valores en el archivo redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Cambiar los valores en el archivo tpch.properties de PrestoSQL.	Not available.
prestosql-connector-tpcds	Cambiar los valores del archivo tpcds.properties de PrestoSQL.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.pro perties.	Restarts Zookeeper server.

Amazon EMR, versión 6.2.0

Versiones de las aplicaciones de la versión 6.2.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Oozie, Phoenix, Pig, PrestoPrestoSQLSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
AWS SDK for Java	1.11.880	1.11,828	1.11.828	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.5	2.2.5	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1
Hue	4.8.0	4.7.1	4.7.1	4.4.0

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	-
Presto	0,238,3	0.232	0.232	0.230
Spark	3.0.1	3.0.0	3.0.0	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	-
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	338	338	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.2.0

Las siguientes notas de la versión incluyen información sobre la versión 6.2.0 de Amazon EMR. Los cambios son respecto a la versión 6.1.0.

Fecha de lanzamiento inicial: 9 de diciembre de 2020

Fecha de la última actualización: 4 de octubre de 2021

Aplicaciones compatibles

- AWS SDK para Java versión 1.11.828
- emr-record-server versión 1.7.0
- Flink, versión 1.11.2
- Versión de Ganglia 3.7.2
- Hadoop, versión 3.2.1-amzn-1
- HBase versión 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog versión 3.1.2-amzn-0
- Hive, versión 3.1.2-amzn-3
- Hudi versión 0.6.0-amzn-1
- Hue versión 4.8.0
- JupyterHub versión 1.1.0
- Livy versión 0.7.0
- MXNet versión 1.7.0
- Oozie versión 5.2.0
- Versión de Phoenix 5.0.0
- Pig, versión 0.17.0
- Presto, versión 0.238.3-amzn-1
- PrestoSQL, versión 343
- Spark, versión 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow versión 2.3.1
- Zeppelin, versión 0.9.0-preview1

- Versión de Zookeeper 3.4.14
- Conectores y controladores: conector de DynamoDB 4.16.0

Nuevas características

- HBase: Se ha eliminado el cambio de nombre en la fase de confirmación y se ha añadido el seguimiento persistente HFile . Consulte el HFile seguimiento persistente en la guía de versiones de Amazon EMR.
- HBase: Backported Cree una configuración que obligue a almacenar en caché los bloques al compactarlos.
- PrestoDB: mejoras en la eliminación dinámica de particiones. El reordenamiento de uniones basado en reglas funciona con datos no particionados.
- Políticas gestionadas con ámbito de aplicación: para ajustarse a las prácticas AWS recomendadas, Amazon EMR ha introducido la versión 2 de las políticas gestionadas predeterminadas con ámbito de EMR como sustitutivas de las políticas que quedarán en desuso. Consulte Políticas administradas por Amazon EMR.
- Estado de soporte de Instance Metadata Service (IMDS) V2: para Amazon EMR 6.2 o posterior, los componentes de Amazon EMR se IMDSv2 utilizan para todas las llamadas de IMDS. Para las llamadas de IMDS en el código de su aplicación, puede utilizar ambos IMDSv1 o configurar el IMDS para que se utilice únicamente con fines de seguridad adicionales. IMDSv2 IMDSv2 Si la deshabilita IMDSv1 en versiones anteriores de Amazon EMR 6.x, se produce un error al iniciar el clúster.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.

- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Spark: mejoras de rendimiento en el tiempo de ejecución de Spark.

Problemas conocidos

- Amazon EMR 6.2 tiene permisos incorrectos configurados en el directorio/etc/cron.d/libinstance-controller-java file in EMR 6.2.0. Permissions on the file are 645 (-rw-r--r-x), when they should be 644 (-rw-r--r--). As a result, Amazon EMR version 6.2 does not log instance-state logs, and the /emr/instance-logs que está vacío. Este problema se ha corregido en la versión 6.3.0 de Amazon EMR y posteriores.

  Para solucionar este problema, ejecute el siguiente script como acción de arranque al iniciar el clúster.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- En el caso de los clústeres de subredes privadas de Amazon EMR 6.2.0 y 6.3.0, no puede acceder a la interfaz de usuario web de Ganglia. Aparecerá el error “Acceso denegado (403)”. Otras

páginas web UIs, como Spark, Hue, Zeppelin JupyterHub, Livy y Tez, funcionan con normalidad. El acceso a la interfaz de usuario web de Ganglia en los clústeres de subredes públicas también funciona con normalidad. Para resolver este problema, reinicie el servicio httpd en el nodo principal con `sudo systemctl restart httpd`. Este problema se ha corregido en la versión 6.4.0 de Amazon EMR.

- Hay un problema en Amazon EMR 6.2.0 por el que httpd genera errores continuamente, lo que provoca que Ganglia no esté disponible. Aparece el error “No se puede conectar al servidor”. Para corregir un clúster que ya se está ejecutando con este problema, utilice SSH en el nodo principal del clúster y agregue la línea `Listen 80` al archivo `httpd.conf` ubicado en `/etc/httpd/conf/httpd.conf`. Este problema se ha corregido en la versión 6.3.0 de Amazon EMR.
- HTTPD presenta errores en los clústeres de EMR 6.2.0 cuando se utiliza una configuración de seguridad. Esto hace que la interfaz de usuario de la aplicación web de Ganglia no esté disponible. Para acceder a la interfaz de usuario de la aplicación web de Ganglia, agregue `Listen 80` al archivo `/etc/httpd/conf/httpd.conf` del nodo principal del clúster. Para obtener información acerca de la conexión al clúster, consulte Conectarse al nodo principal mediante SSH.

  Cuadernos de EMR tampoco puede establecer una conexión con los clústeres de EMR 6.2.0 cuando se utiliza una configuración de seguridad. El cuaderno no podrá enumerar los kernels ni enviar trabajos a Spark. En su lugar, se recomienda que utilice Cuadernos de EMR con otra versión de Amazon EMR.
- Reducir el límite de «máximo de archivos abiertos» en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
```

```
sudo systemctl daemon-reload
```

- > **Important**
  >
  > Amazon EMR 6.1.0 y 6.2.0 incluyen un problema de rendimiento que puede afectar gravemente a todas las operaciones de inserción, upsert y eliminación de Hudi. Si piensa utilizar Hudi con Amazon EMR 6.1.0 o 6.2.0, póngase en AWS contacto con el servicio de asistencia para obtener un RPM de Hudi parcheado.

- > **Important**
  >
  > Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan automáticamente esas actualizaciones. Para obtener actualizaciones del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.

- Los artefactos de Maven de Amazon EMR 6.2.0 no están publicados. Se publicarán en una versión futura de Amazon EMR.
- El HFile seguimiento persistente mediante la tabla del sistema de HBase archivos de almacenamiento no admite la función de replicación HBase regional. Para obtener más información sobre HBase la replicación regional, consulte Lecturas de alta disponibilidad consistentes en el cronograma.
- Diferencias entre las versiones de agrupación en buckets de Amazon EMR 6.x y EMR 5.x Hive

  EMR 5.x utiliza OOS Apache Hive 2, mientras que en EMR 6.x se utiliza OOS Apache Hive 3. Hive2 de código abierto usa la versión 1 de agrupación en buckets, mientras que Hive3 de código abierto usa la versión 2 de agrupación en buckets. Esta diferencia de versión de agrupación en buckets entre Hive 2 (EMR 5.x) y Hive 3 (EMR 6.x) significa que el hash de agrupación en buckets de Hive funciona de manera diferente. Vea el ejemplo siguiente.

  La siguiente tabla es un ejemplo creado en EMR 6.x y EMR 5.x, respectivamente.

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

Se insertan los mismos datos en EMR 6.x y EMR 5.x.

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

Al comprobar la ubicación de S3, se muestra que el nombre del archivo de agrupación en buckets es diferente, ya que la función de hash es diferente entre EMR 6.x (Hive 3) y EMR 5.x (Hive 2).

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

También puede ver la diferencia de versión mediante la ejecución del siguiente comando en la CLI de Hive en EMR 6.x. Tenga en cuenta que devuelve la versión 2 de la agrupación en buckets.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del

periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 6.2.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-notebook-env	1.0.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.16.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.0.0	Conector S3Select de EMR
emrfs	2.44.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.11.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-2	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	3.2.1-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.2.6-amzn-0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.2.6-amzn-0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.6-amzn-0	HBase cliente de línea de comandos.
hbase-rest-server	2.2.6-amzn-0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.6-amzn-0	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	3.1.2-amzn-3	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-3	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-3	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-3	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-3	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-3	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-3	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.6.0-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.6.0-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hudi-prestosql	0.6.0-amzn-1	Biblioteca de paquetes para ejecutar PrestoSQL con Hudi.
hudi-spark	0.6.0-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.8.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64+	Servidor de base de datos de MariaDB.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.

Componente	Versión	Descripción
opencv	4.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.0.0- -2.0 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.238.3-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.238.3-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.238.3-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
prestosql-coordinator	343	Servicio para aceptar las consultas y administrar la ejecución de consultas entre prestosql-workers.
prestosql-worker	343	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
prestosql-client	343	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.0.1-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.0.1-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.0.1-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.0.1-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
spark-rapids	0.2.0	El complemento Nvidia Spark RAPIDS que acelera Apache Spark con... GPUs
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.3.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0-preview1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.2.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-6.2.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Not available.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Not available.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	This classification should not be reconfigured.
hdfs-env	Cambiar los valores en el entorno de HDFS.	Restarts Hadoop HDFS ZKFC.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat server.
hive	Configuración de Amazon EMR mantenida para Apache Hive.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Not available.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoDB)
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
prestosql-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server (for PrestoSQL)

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	Cambiar los valores en el archivo node.properties de PrestoSQL.	Not available.
prestosql-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de PrestoSQL.	Not available.
prestosql-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de PrestoSQL.	Not available.
prestosql-connector-hive	Cambiar los valores en el archivo hive.properties de PrestoSQL.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	Cambiar los valores en el archivo jmx.properties de PrestoSQL.	Not available.
prestosql-connector-kafka	Cambiar los valores en el archivo kafka.properties de PrestoSQL.	Not available.
prestosql-connector-localfile	Cambiar los valores en el archivo localfile.properties de PrestoSQL.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
prestosql-connector-memory	Cambiar los valores del archivo memory.properties de PrestoSQL.	Not available.
prestosql-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de PrestoSQL.	Not available.
prestosql-connector-mysql	Cambiar los valores en el archivo mysql.properties de PrestoSQL.	Not available.
prestosql-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de PrestoSQL.	Not available.
prestosql-connector-raptor	Cambiar los valores en el archivo raptor.properties de PrestoSQL.	Not available.
prestosql-connector-redis	Cambiar los valores en el archivo redis.properties de PrestoSQL.	Not available.
prestosql-connector-redshift	Cambiar los valores en el archivo redshift.properties de PrestoSQL.	Not available.
prestosql-connector-tpch	Cambiar los valores en el archivo tpch.properties de PrestoSQL.	Not available.
prestosql-connector-tpcds	Cambiar los valores del archivo tpcds.properties de PrestoSQL.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restart Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 6.1.1

Versiones de las aplicaciones de la versión 6.1.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, JupyterHub, Livy, MXNet,, Oozie, Phoenix, Pig, Presto, PrestoSQLSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.1.1

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

Cambios, mejoras y problemas resueltos

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.

- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de AWS debate. Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Versiones de los componentes de la versión 6.1.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la

comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.3.0	SDK de Amazon SageMaker Spark
emr-ddb	4.14.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.14.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.0.0	Conector S3Select de EMR
emrfs	2.42.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-1.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-1.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-1.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-1.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-1.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-1.1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	3.2.1-amzn-1.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-1.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-1.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-1.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-1.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.2.5	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.2.5	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.5	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	2.2.5	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.5	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-2	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hudi	0.5.2-incubating-amzn-2	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.2-incubating-amzn-2	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-prestosql	0.5.2-incubating-amzn-2	Biblioteca de paquetes para ejecutar PrestoSQL con Hudi.
hudi-spark	0.5.2-incubating-amzn-2	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.7.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.6.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64+	Servidor de base de datos de MariaDB.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.3.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.0.0- -2.0 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.232	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.232	Service para ejecutar partes de una consulta.
presto-client	0.232	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.

Componente	Versión	Descripción
prestosql-coordinator	338	Servicio para aceptar las consultas y administrar la ejecución de consultas entre prestosql-workers.
prestosql-worker	338	Service para ejecutar partes de una consulta.
prestosql-client	338	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.0.0-amzn-0.1	Clientes de línea de comando de Spark.
spark-history-server	3.0.0-amzn-0.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.0.0-amzn-0.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	3.0.0-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.1.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0-preview1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.1.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-6.1.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop

Clasificaciones	Descripción
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-env	Cambiar los valores en el entorno de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.

Clasificaciones	Descripción
hive	Configuración de Amazon EMR mantenida para Apache Hive.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.

Clasificaciones	Descripción
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.

Clasificaciones	Descripción
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.

Clasificaciones	Descripción
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
prestosql-log	Cambiar los valores en el archivo log.properties de Presto.
prestosql-config	Cambiar los valores en el archivo config.properties de Presto.
prestosql-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
prestosql-env	Cambiar valores en el archivo presto-env.sh de Presto.
prestosql-node	Cambiar los valores en el archivo node.properties de PrestoSQL.
prestosql-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de PrestoSQL.
prestosql-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de PrestoSQL.
prestosql-connector-hive	Cambiar los valores en el archivo hive.properties de PrestoSQL.
prestosql-connector-jmx	Cambiar los valores en el archivo jmx.properties de PrestoSQL.
prestosql-connector-kafka	Cambiar los valores en el archivo kafka.properties de PrestoSQL.
prestosql-connector-localfile	Cambiar los valores en el archivo localfile.properties de PrestoSQL.

Clasificaciones	Descripción
prestosql-connector-memory	Cambiar los valores del archivo memory.properties de PrestoSQL.
prestosql-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de PrestoSQL.
prestosql-connector-mysql	Cambiar los valores en el archivo mysql.properties de PrestoSQL.
prestosql-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de PrestoSQL.
prestosql-connector-raptor	Cambiar los valores en el archivo raptor.properties de PrestoSQL.
prestosql-connector-redis	Cambiar los valores en el archivo redis.properties de PrestoSQL.
prestosql-connector-redshift	Cambiar los valores en el archivo redshift.properties de PrestoSQL.
prestosql-connector-tpch	Cambiar los valores en el archivo tpch.properties de PrestoSQL.
prestosql-connector-tpcds	Cambiar los valores del archivo tpcds.properties de PrestoSQL.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.

Clasificaciones	Descripción
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Clasificaciones	Descripción
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 6.1.0

Versiones de las aplicaciones de la versión 6.1.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hudi, Hue, JupyterHub, Livy, MXNet,, Oozie, Phoenix, Pig, Presto, PrestoSQLSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.1.0

Las siguientes notas de la versión incluyen información sobre la versión 6.1.0 de Amazon EMR. Los cambios son respecto a la versión 6.0.0.

Fecha de lanzamiento inicial: 4 de septiembre de 2020

Fecha de la última actualización: 15 de octubre de 2020

Aplicaciones compatibles

- AWS SDK para Java versión 1.11.828
- Flink, versión 1.11.0
- Versión de Ganglia 3.7.2
- Hadoop, versión 3.2.1-amzn-1
- HBase versión 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog versión 3.1.2-amzn-0
- Hive, versión 3.1.2-amzn-1

- Hudi versión 0.5.2-incubating
- Hue versión 4.7.1
- JupyterHub versión 1.1.0
- Livy versión 0.7.0
- MXNet versión 1.6.0
- Oozie versión 5.2.0
- Versión de Phoenix 5.0.0
- Presto, versión 0.232
- PrestoSQL, versión 338
- Spark, versión 3.0.0-amzn-0
- TensorFlow versión 2.1.0
- Zeppelin, versión 0.9.0-preview1
- Versión de Zookeeper 3.4.14
- Conectores y controladores: DynamoDB Connector 4.14.0

Nuevas características

- Los tipos de instancias de ARM se admiten a partir de la versión 5.30.0 de Amazon EMR y la versión 6.1.0 de Amazon EMR.
- Los tipos de instancia de uso general m6g se admiten a partir de las versiones 6.1.0 y 5.30.0 de Amazon EMR. Para más información, consulte Tipos de instancias admitidas en la Guía de administración de Amazon EMR.
- La función de grupo de EC2 ubicación se admite a partir de la versión 5.23.0 de Amazon EMR como opción para varios clústeres de nodos principales. Actualmente, la característica de grupos de ubicación solo admite los tipos de nodos principales y la estrategia `SPREAD` se aplica a esos nodos. La estrategia `SPREAD` coloca un pequeño grupo de instancias en un hardware subyacente independiente para evitar la pérdida de varios nodos principales en caso de que se produzca un error de hardware. Para obtener más información, consulte Integración de EMR con EC2 Placement Group en la Guía de administración de Amazon EMR.
- Escalado administrado: con la versión 6.1.0 de Amazon EMR, puede habilitar el escalado administrado de Amazon EMR para aumentar o disminuir automáticamente el número de instancias o unidades del clúster en función de la carga de trabajo. Amazon EMR evalúa continuamente las métricas del clúster para tomar decisiones de escalado que optimicen los

clústeres en cuanto al costo y la velocidad. El escalado administrado también está disponible en la versión 5.30.0 y posteriores de Amazon EMR, excepto 6.0.0. Para más información, consulte Recursos de escalado de clúster en la Guía de administración de Amazon EMR.

- La versión 338 de PrestoSQL es compatible con EMR 6.1.0. Para más información, consulte Presto.
  - PrestoSQL solo es compatible con EMR 6.1.0 y versiones posteriores, no con EMR 6.0.0 ni EMR 5.x.
  - El nombre de la aplicación, `Presto`, se sigue utilizando para instalar PrestoDB en los clústeres. Para instalar PrestoSQL en los clústeres, utilice el nombre de la aplicación `PrestoSQL`.
  - Puede instalar PrestoDB o PrestoSQL, pero no puede instalar ambos en un solo clúster. Si se especifican PrestoDB y PrestoSQL al intentar crear un clúster, se produce un error tanto en la validación como en la solicitud de creación del clúster.
  - PrestoSQL es compatible con clústeres de maestro único o multimaestro. En los clústeres multimaestro, se necesita un metaalmacén de Hive externo para ejecutar PrestoSQL o PrestoDB. Consulte Aplicaciones admitidas en un clúster de EMR con varios nodos principales.
- Soporte de autenticación automática de ECR en Apache Hadoop y Apache Spark con Docker: los usuarios de Spark pueden utilizar imágenes de Docker de Docker Hub y Amazon Elastic Container Registry (Amazon ECR) para definir las dependencias del entorno y la biblioteca.

  Configuración de Docker y Ejecución de aplicaciones de Spark con Docker mediante Amazon EMR 6.x..
- EMR admite transacciones ACID de Apache Hive: Amazon EMR 6.1.0 agrega compatibilidad con las transacciones ACID de Hive para que cumpla con las propiedades ACID de una base de datos. Con esta característica, puede ejecutar operaciones `INSERT, UPDATE, DELETE,` y `MERGE` en tablas administradas por Hive con datos de Amazon Simple Storage Service (Amazon S3). Se trata de una característica clave para casos de uso como la ingesta de contenido en streaming, la reformulación de datos, las actualizaciones masivas mediante MERGE y los cambios lentos de dimensiones. Para más información, incluidos ejemplos de configuración y casos de uso, consulte Amazon EMR admite transacciones ACID de Apache Hive.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Apache Flink no es compatible con EMR 6.0.0, pero sí con EMR 6.1.0 con Flink 1.11.0. Esta es la primera versión de Flink que es compatible oficialmente con Hadoop 3. Consulte Apache Flink 1.11.0 Release Announcement.
- Se ha eliminado Ganglia de los paquetes predeterminados de EMR 6.1.0.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0,

emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
```

```
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon EMR 6.1.0 y 6.2.0 incluyen un problema de rendimiento que puede afectar gravemente a todas las operaciones de inserción, upsert y eliminación de Hudi. Si piensa utilizar Hudi con Amazon EMR 6.1.0 o 6.2.0, póngase en AWS contacto con el servicio de asistencia para obtener un RPM de Hudi parcheado.

- Si establece una configuración de recolección de basura personalizada con `spark.driver.extraJavaOptions` y`spark.executor.extraJavaOptions`, esto provocará un error de driver/executor inicio con EMR 6.1 debido a un conflicto en la configuración de recolección de basura. En su lugar, con la versión 6.1.0 de EMR, debe especificar una configuración personalizada de recopilación de elementos no utilizados de Spark para los controladores y ejecutores con las propiedades `spark.driver.defaultJavaOptions` y `spark.executor.defaultJavaOptions`. Obtenga más información en Apache Spark Runtime Environment y Configuración de la recopilación de elementos no utilizados de Spark en Amazon EMR 6.1.0.
- Al utilizar Pig con Oozie (y dentro de Hue, ya que Hue utiliza las acciones de Oozie para ejecutar scripts de Pig), se genera un error que indica que no se puede cargar una biblioteca lzo nativa. Este mensaje de error es informativo y no impide que Pig se ejecute.
- Soporte de simultaneidad de Hudi: actualmente, Hudi no admite escrituras simultáneas en una sola tabla de Hudi. Además, Hudi revierte los cambios que estén realizando los escritores en curso antes de permitir que comience un nuevo escritor. Las escrituras simultáneas pueden interferir con

este mecanismo e introducir condiciones de carrera, lo que puede provocar la corrupción de los datos. Debe asegurarse de que, como parte de su flujo de trabajo de procesamiento de datos, solo haya un escritor de Hudi trabajando en una tabla de Hudi a la vez. Hudi admite que varios lectores simultáneos trabajen en la misma tabla de Hudi.

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Hay un problema en Amazon EMR 6.1.0 que afecta a los clústeres que ejecutan Presto. Tras un periodo de tiempo prolongado (días), el clúster puede generar errores como “su: no se puede ejecutar /bin/bash: recurso temporalmente no disponible” o “error de la solicitud del intérprete de comandos en el canal 0”. Este problema se debe a un proceso interno de Amazon EMR (InstanceController) que genera demasiados procesos ligeros (LWP), lo que, finalmente, hace

que el usuario de Hadoop supere su límite de nproc. Esto impide que el usuario abra procesos adicionales. La solución para este problema consiste en actualizar a EMR 6.2.0.

Versiones de los componentes de la versión 6.1.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.3.0	SDK de Amazon SageMaker Spark
emr-ddb	4.14.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.14.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	2.0.0	Conector S3Select de EMR
emrfs	2.42.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	3.2.1-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	3.2.1-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.2.5	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.2.5	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.5	HBase cliente de línea de comandos.
hbase-rest-server	2.2.5	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.5	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	3.1.2-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-2	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.2-incubating-amzn-2	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.2-incubating-amzn-2	Biblioteca de paquetes para ejecutar Presto con Hudi.
hudi-prestosql	0.5.2-incubating-amzn-2	Biblioteca de paquetes para ejecutar PrestoSQL con Hudi.

Componente	Versión	Descripción
hudi-spark	0.5.2-incubating-amzn-2	Biblioteca de paquetes para ejecutar Spark con Hudi.
hue-server	4.7.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.6.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64+	Servidor de base de datos de MariaDB.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.3.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	5.0.0- -2.0 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.232	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.232	Service para ejecutar partes de una consulta.
presto-client	0.232	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
prestosql-coordinator	338	Servicio para aceptar las consultas y administrar la ejecución de consultas entre prestosql-workers.
prestosql-worker	338	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
prestosql-client	338	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	2.0.0	Sistema de administración de claves Apache Ranger
spark-client	3.0.0-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	3.0.0-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	3.0.0-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	3.0.0-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.1.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0-preview1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.1.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-6.1.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.

Clasificaciones	Descripción
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-env	Cambiar los valores en el entorno de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive	Configuración de Amazon EMR mantenida para Apache Hive.

Clasificaciones	Descripción
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.

Clasificaciones	Descripción
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
prestosql-log	Cambiar los valores en el archivo log.properties de Presto.
prestosql-config	Cambiar los valores en el archivo config.properties de Presto.
prestosql-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
prestosql-env	Cambiar valores en el archivo presto-env.sh de Presto.
prestosql-node	Cambiar los valores en el archivo node.properties de PrestoSQL.
prestosql-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de PrestoSQL.
prestosql-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de PrestoSQL.
prestosql-connector-hive	Cambiar los valores en el archivo hive.properties de PrestoSQL.
prestosql-connector-jmx	Cambiar los valores en el archivo jmx.properties de PrestoSQL.
prestosql-connector-kafka	Cambiar los valores en el archivo kafka.properties de PrestoSQL.
prestosql-connector-localfile	Cambiar los valores en el archivo localfile.properties de PrestoSQL.
prestosql-connector-memory	Cambiar los valores del archivo memory.properties de PrestoSQL.

Clasificaciones	Descripción
prestosql-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de PrestoSQL.
prestosql-connector-mysql	Cambiar los valores en el archivo mysql.properties de PrestoSQL.
prestosql-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de PrestoSQL.
prestosql-connector-raptor	Cambiar los valores en el archivo raptor.properties de PrestoSQL.
prestosql-connector-redis	Cambiar los valores en el archivo redis.properties de PrestoSQL.
prestosql-connector-redshift	Cambiar los valores en el archivo redshift.properties de PrestoSQL.
prestosql-connector-tpch	Cambiar los valores en el archivo tpch.properties de PrestoSQL.
prestosql-connector-tpcds	Cambiar los valores del archivo tpcds.properties de PrestoSQL.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop. OraOop
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 6.0.1

Versiones de las aplicaciones de la versión 6.0.1

Esta versión incluye las siguientes aplicaciones: Ganglia,, HBase,, HCatalog, Hadoop, Hive, Hudi, Hue,, JupyterHub, Livy, MXNet, Oozie, Phoenix, Presto,, SparkTensorFlowTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.0.1

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

Cambios, mejoras y problemas resueltos

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.

- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de AWS debate. Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Versiones de los componentes de la versión 6.0.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.6	SDK de Amazon SageMaker Spark
emr-ddb	4.14.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.14.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.5.0	Conector S3Select de EMR
emrfs	2.39.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-0.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-0.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-0.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-0.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	3.2.1-amzn-0.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-0.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-0.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	3.2.1-amzn-0.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-0.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-0.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-0.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	2.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.3	HBase cliente de línea de comandos.
hbase-rest-server	2.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.3	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	3.1.2-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-0	Hive-hbase client.
hive-metastore-server	3.1.2-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.0-incubating-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.0-incubating-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.0.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64+	Servidor de base de datos de MariaDB.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.0.0- HBase -2.0	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.230	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.230	Service para ejecutar partes de una consulta.
presto-client	0.230	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
r	3.4.3	Proyecto R para análisis estadístico
spark-client	2.4.4	Clientes de línea de comando de Spark.
spark-history-server	2.4.4	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.4	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.4.4	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.
zeppelin-server	0.9.0-SNAPSHOT	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.0.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-6.0.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase

Clasificaciones	Descripción
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-env	Cambiar los valores en el entorno de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive	Configuración de Amazon EMR mantenida para Apache Hive.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.

Clasificaciones	Descripción
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.

Clasificaciones	Descripción
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.

Clasificaciones	Descripción
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Clasificaciones	Descripción
zookeeper-config	Cambie los valores en ZooKeeper el archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 6.0.0

Versiones de las aplicaciones de la versión 6.0.0

Esta versión incluye las siguientes aplicaciones: Ganglia,, HBase,, HCatalog, Hadoop, Hive, Hudi, Hue,, JupyterHub, Livy, MXNet, Oozie, Phoenix, Presto,, SparkTensorFlowTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS SDK for Java	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 6.0.0

Las siguientes notas de la versión incluyen información sobre la versión 6.0.0 de Amazon EMR.

Fecha de lanzamiento inicial: 10 de marzo de 2020

Aplicaciones compatibles

- AWS SDK para Java versión 1.11.711
- Versión de Ganglia 3.7.2
- Versión de Hadoop 3.2.1
- HBase versión 2.2.3
- HCatalog versión 3.1.2
- Versión de Hive 3.1.2
- Versión de Hudi 0.5.0-incubating
- Versión de Hue 4.4.0
- JupyterHub versión 1.0.0
- Versión de Livy 0.6.0

- MXNet versión 1.5.1
- Versión de Oozie 5.1.0
- Versión de Phoenix 5.0.0
- Versión de Presto 0.230
- Versión de Spark 2.4.4
- TensorFlow versión 1.14.0
- Versión de Zeppelin 0.9.0-SNAPSHOT
- Versión de Zookeeper 3.4.14
- Conectores y controladores: DynamoDB Connector 4.14.0

> **Note**
>
> Flink, Sqoop, Pig y Mahout no están disponibles en la versión 6.0.0 de Amazon EMR.

Nuevas características

- Soporte de tiempo de ejecución para YARN Docker: las aplicaciones de YARN, como los trabajos de Spark, ahora pueden ejecutarse en el contexto de un contenedor de Docker. Esto le permite definir fácilmente dependencias en una imagen de Docker sin necesidad de instalar bibliotecas personalizadas en su clúster de Amazon EMR. Para más información, consulte Configurar la integración de Docker y Ejecutar aplicaciones de Spark con Docker mediante Amazon EMR 6.0.0.
- Soporte LLAP de Hive: Hive ahora admite el modo de ejecución LLAP para mejorar el rendimiento de las consultas. Para obtener más información, consulte Uso de Hive LLAP.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.

- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Amazon Linux
  - Amazon Linux 2 es el sistema operativo de la serie de versiones 6.x de EMR.
  - Para la administración de servicios se utiliza `systemd` en lugar de `upstart` que se utilizaba en Amazon Linux 1.
- Kit de desarrollo de Java (JDK)
  - Corretto JDK 8 es el JDK predeterminado para la serie de versiones de EMR 6.x.
- Scala
  - Scala 2.12 se utiliza con Apache Spark y Apache Livy.
- Python 3
  - Python 3 es ahora la versión predeterminada de Python en EMR.
- Etiquetas de nodo YARN

- A partir de la serie de versiones 6.x de Amazon EMR, la característica de etiquetas de nodo YARN está desactivada de forma predeterminada. De forma predeterminada, los procesos maestros de la aplicación se pueden ejecutar tanto en nodos principales como en nodos de tareas. Puede habilitar la función de etiquetas de nodo YARN configurando las siguientes propiedades: `yarn.node-labels.enabled` y `yarn.node-labels.am.default-node-label-expression`. Para más información, consulte Comprensión de los nodos principales, básicos y de tareas.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

  Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

  Establecimiento explícito de un ulimit desde la línea de comandos

  1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

2. Reinicie InstanceController

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- El shell interactivo de Spark, que incluye PySpark SparkR y spark-shell, no admite el uso de Docker con bibliotecas adicionales.
- Para utilizar Python 3 con la versión 6.0.0 de Amazon EMR, debe agregar `PATH` a `yarn.nodemanager.env-whitelist`.
- La funcionalidad Live Long and Process (LLAP) no se admite cuando se utiliza el catálogo de datos de AWS Glue como metaalmacén de Hive.
- Al utilizar Amazon EMR 6.0.0 con la integración de Spark y Docker, debe configurar las instancias del clúster con el mismo tipo de instancia y la misma cantidad de volúmenes de EBS para evitar errores al enviar un trabajo de Spark con el tiempo de ejecución de Docker.

- En Amazon EMR 6.0.0, el modo de almacenamiento de Amazon HBase S3 se ve afectado por el problema del HBASE-24286. HBase el maestro no se puede inicializar cuando se crea el clúster con los datos de S3 existentes.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 6.0.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.6	SDK de Amazon SageMaker Spark
emr-ddb	4.14.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	3.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.14.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.5.0	Conector S3Select de EMR
emrfs	2.39.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	3.2.1-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	3.2.1-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	3.2.1-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	3.2.1-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	3.2.1-amzn-0	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	3.2.1-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	3.2.1-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	3.2.1-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	3.2.1-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	3.2.1-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	3.2.1-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	2.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	2.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	2.2.3	HBase cliente de línea de comandos.
hbase-rest-server	2.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	2.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	3.1.2-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	3.1.2-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	3.1.2-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	3.1.2-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	3.1.2-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	3.1.2-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	3.1.2-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.0-incubating-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.0-incubating-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.0.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.

Componente	Versión	Descripción
mariadb-server	5.5.64+	Servidor de base de datos de MariaDB.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	5.0.0- HBase -2.0	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	5.0.0- -2.0 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.230	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.230	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.230	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
r	3.4.3	Proyecto R para análisis estadístico
spark-client	2.4.4	Clientes de línea de comando de Spark.
spark-history-server	2.4.4	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.4	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.4	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.41+	Servidor HTTP de Apache.

Componente	Versión	Descripción
zeppelin-server	0.9.0-SNAPSHOT	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 6.0.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-6.0.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.

Clasificaciones	Descripción
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-env	Cambiar los valores en el entorno de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.

Clasificaciones	Descripción
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive	Configuración de Amazon EMR mantenida para Apache Hive.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.

Clasificaciones	Descripción
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.

Clasificaciones	Descripción
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.

Clasificaciones	Descripción
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.pr operties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.

Clasificaciones	Descripción
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en ZooKeeper el archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Versiones de lanzamiento de Amazon EMR 5.x

Esta sección contiene las versiones de las aplicaciones, las notas de la versión, las versiones de los componentes y las clasificaciones de configuración de cada versión de lanzamiento de Amazon EMR 5.x.

Al lanzar un clúster, puede elegir entre varias versiones de Amazon EMR. Esto le permite probar y utilizar versiones de la aplicación que cumplan con sus requisitos de compatibilidad. El número de versión se especifica con la etiqueta de versión. Las etiquetas de versión tienen el formato `emr-x.x.x` Por ejemplo, `emr-7.9.0`.

Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Para obtener una tabla completa de las versiones de las aplicaciones incluidas en cada versión de Amazon EMR 5.x, consulte Versiones de la aplicación en las versiones 5.x de Amazon EMR.

Temas

- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Amazon EMR, versión 5.36.2
- Amazon EMR, versión 5.36.1
- Amazon EMR, versión 5.36.0
- Amazon EMR, versión 5.35.0
- Amazon EMR, versión 5.34.0
- Amazon EMR, versión 5.33.1
- Amazon EMR, versión 5.33.0
- Amazon EMR, versión 5.32.1
- Amazon EMR, versión 5.32.0
- Amazon EMR, versión 5.31.1
- Amazon EMR, versión 5.31.0
- Amazon EMR, versión 5.30.2
- Amazon EMR, versión 5.30.1
- Amazon EMR, versión 5.30.0
- Amazon EMR, versión 5.29.0
- Amazon EMR, versión 5.28.1
- Amazon EMR, versión 5.28.0
- Amazon EMR, versión 5.27.1
- Amazon EMR, versión 5.27.0
- Amazon EMR, versión 5.26.0
- Amazon EMR, versión 5.25.0

- Amazon EMR, versión 5.24.1
- Amazon EMR, versión 5.24.0
- Amazon EMR, versión 5.23.1
- Amazon EMR, versión 5.23.0
- Amazon EMR, versión 5.22.0
- Amazon EMR, versión 5.21.2
- Amazon EMR, versión 5.21.1
- Amazon EMR, versión 5.21.0
- Amazon EMR, versión 5.20.1
- Amazon EMR, versión 5.20.0
- Amazon EMR, versión 5.19.1
- Amazon EMR, versión 5.19.0
- Amazon EMR, versión 5.18.1
- Amazon EMR, versión 5.18.0
- Amazon EMR, versión 5.17.2
- Amazon EMR, versión 5.17.1
- Amazon EMR, versión 5.17.0
- Amazon EMR, versión 5.16.1
- Amazon EMR, versión 5.16.0
- Amazon EMR, versión 5.15.1
- Amazon EMR, versión 5.15.0
- Amazon EMR, versión 5.14.2
- Amazon EMR, versión 5.14.1
- Amazon EMR, versión 5.14.0
- Amazon EMR, versión 5.13.1
- Amazon EMR, versión 5.13.0
- Amazon EMR, versión 5.12.3
- Amazon EMR, versión 5.12.2

- Amazon EMR, versión 5.12.1
- Amazon EMR, versión 5.12.0
- Amazon EMR, versión 5.11.4
- Amazon EMR, versión 5.11.3
- Amazon EMR, versión 5.11.2
- Amazon EMR, versión 5.11.1
- Amazon EMR, versión 5.11.0
- Amazon EMR, versión 5.10.1
- Amazon EMR, versión 5.10.0
- Amazon EMR, versión 5.9.1
- Amazon EMR, versión 5.9.0
- Amazon EMR, versión 5.8.3
- Amazon EMR, versión 5.8.2
- Amazon EMR, versión 5.8.1
- Amazon EMR, versión 5.8.0
- Amazon EMR, versión 5.7.1
- Amazon EMR, versión 5.7.0
- Amazon EMR, versión 5.6.1
- Amazon EMR, versión 5.6.0
- Amazon EMR, versión 5.5.4
- Amazon EMR, versión 5.5.3
- Amazon EMR, versión 5.5.2
- Amazon EMR, versión 5.5.1
- Amazon EMR, versión 5.5.0
- Amazon EMR, versión 5.4.1
- Amazon EMR, versión 5.4.0
- Amazon EMR, versión 5.3.2
- Amazon EMR, versión 5.3.1

- Amazon EMR, versión 5.3.0
- Amazon EMR, versión 5.2.3
- Amazon EMR, versión 5.2.2
- Amazon EMR, versión 5.2.1
- Amazon EMR, versión 5.2.0
- Amazon EMR, versión 5.1.1
- Amazon EMR, versión 5.1.0
- Amazon EMR, versión 5.0.3
- Amazon EMR, versión 5.0.2
- Amazon EMR, versión 5.0.1
- Amazon EMR, versión 5.0.0

Versiones de la aplicación en las versiones 5.x de Amazon EMR

Para ver una tabla completa con las versiones de las aplicaciones disponibles en cada versión de Amazon EMR 5.x, abra Versiones de la aplicación en las versiones 5.x de Amazon EMR en su navegador.

Amazon EMR, versión 5.36.2

Versiones de las aplicaciones de la versión 5.36.2

Esta versión incluye las siguientes aplicaciones: Delta, Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, Mahout, Oozie, PhoenixPig Presto, Spark, Sqoop, TensorFlowTez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR

- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
AWS SDK for Java	1.12.206	1.12.206	1.12.206	1.12.159
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.9
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.9
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,267	0,267	0,266
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.36.2

Las siguientes notas de la versión incluyen información sobre la versión 5.36.2 de Amazon EMR. Los cambios son respecto a la versión 5.36.1. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Cambios, mejoras y problemas resueltos

- Esta versión mejora la lógica de reducción vertical del clúster para que Amazon EMR no intente reducir verticalmente los nodos básicos por debajo de la configuración del factor de replicación

de HDFS del clúster. Esta mejora cumple con los requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.

- Esta versión agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se ejecute indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.
- Se corrige un problema en el que las operaciones de reducción de escala del clúster pueden detenerse mientras Amazon EMR desactiva un nodo básico de forma gradual y este pasa a mal estado antes de desactivarse por completo..
- Se mejora la estabilidad de un nodo en un clúster de alta disponibilidad con varios nodos básicos cuando Amazon EMR reinicia un único nodo.
- Optimiza la administración de registros con Amazon EMR que se ejecuta en Amazon. EC2 Como resultado, es posible que observe una pequeña reducción en los costos de almacenamiento de los registros de su clúster.
- Mejora la administración de los archivos de registro de ZooKeeper transacciones que se mantienen en los nodos principales para minimizar los escenarios en los que los archivos de registro se exceden de los límites e interrumpen las operaciones del clúster.
- Corrige un error poco frecuente que podía provocar el fallo de un clúster de alta disponibilidad con varios nodos principales por no poder comunicarse con el Yarn. ResourceManager
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyu Derabad), EE.UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,14,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 503.0	4,1343	xxxxxx, 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

Versiones de los componentes de la versión 5.36.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.16.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.5.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter

Componente	Versión	Descripción
emr-s3-dist-cp	2.21.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.7.0	Conector S3Select de EMR
emrfs	2,51,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.14.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.14.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.10.1-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.10.1-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.10.1-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.9-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.9-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.9-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.9-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.9-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.9-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.9-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.10.1-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.10.1-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.10.1-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.13.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68	Servidor de bases de datos MySQL.
nvidia-cuda	11.0.194	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.267-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.267-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.267-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.8-amzn-2	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.4.8-amzn-2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.8-amzn-2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.8-amzn-2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.36.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.36.2

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Clasificaciones	Descripción	Acciones de reconfiguración
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.

Clasificaciones	Descripción	Acciones de reconfiguración
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 5.36.2

Registro de cambios de la versión 5.36.2 y notas de la versión

Date	Evento	Descripción
28-05-2022	Implementación completa	Amazon EMR 5.36.2 se ha implementado completamente en todas las regiones compatibles
2024-05-28	Publicación de documentos	Las notas de la versión 5.36.2 de Amazon EMR se publicaron por primera vez
2024-05-20	Versión inicial	Amazon EMR 5.36.2 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 5.36.1

Versiones de las aplicaciones de la versión 5.36.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia,, HBase, HCatalog, Hadoop,, Hive, Hudi, Hue, Iceberg,, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Mahout, OoziePhoenixPigPresto Spark, Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
AWS SDK for Java	1.12.206	1.12.206	1.12.159	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.8
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.8
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,267	0,266	0,261
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.36.1

Las siguientes notas de la versión incluyen información sobre la versión 5.36.1 de Amazon EMR. Los cambios son respecto a la versión 5.36.0. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Cambios, mejoras y problemas resueltos

- La versión 5.36.1 de Amazon EMR agrega compatibilidad para archivar registros en Amazon S3 durante la reducción vertical del clúster. En las versiones 5.x anteriores, solo podía archivar archivos de registro en Amazon S3 durante la terminación del clúster. Esta mejora garantiza que los archivos de registro generados en el clúster persistan en Amazon S3 incluso después de terminar el nodo. Para más información, consulte Configurar el registro y la depuración de un clúster.
- La versión 5.36.1 mejora el daemon de administración de registros en el clúster para supervisar las carpetas de registro adicionales del clúster de EMR. Esta mejora minimiza los escenarios de sobreutilización del disco.
- La versión 5.36.1 reinicia automáticamente el daemon de administración de registros en el clúster cuando se detiene. Esta mejora reduce el riesgo de que los nodos parezcan estar en mal estado debido a la sobreutilización del disco.
- La versión 5.36.1 corrige un problema por el que los daemons de Amazon EMR del nodo principal mantenían metadatos obsoletos para las instancias terminadas del clúster. Conservar datos obsoletos puede provocar que el uso de la CPU y la memoria en el clúster aumente sin límites y, en última instancia, causar fallos en el clúster.
- Para los clústeres que se lanzan con varios nodos principales, la versión 5.36.1 corrige un problema por el que un fallo de EC2 hardware de Amazon en uno de los nodos principales podía provocar un error en un segundo nodo principal y hacer que el clúster fuera inestable.

- En el caso de los clústeres que están configurados con cifrado en tránsito, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark. Los datos de mezclas aleatorias de Spark son datos que Spark redistribuye entre las particiones para realizar operaciones específicas. Durante la reducción vertical, el escalado administrado ignora las instancias con datos de mezclas aleatorias. Esto evita que se vuelvan a intentar los trabajos y se vuelvan a realizar los cálculos, lo que resulta costoso en términos de precio y rendimiento. Para más información sobre las operaciones de mezclas aleatorias, consulte Spark Programming Guide.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE. UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,14,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 131,0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 124,0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 218,0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 206,0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 116,0	4,14,328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 101,0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 020,1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 012,1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 906,0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4.14.314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

Versiones de los componentes de la versión 5.36.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.16.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.5.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.21.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.7.0	Conector S3Select de EMR

Componente	Versión	Descripción
emrfs	2,51,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.14.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.14.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.10.1-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.10.1-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.10.1-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.9-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.9-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.9-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.9-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.9-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.9-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.9-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.10.1-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.10.1-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.10.1-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.13.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68	Servidor de bases de datos MySQL.
nvidia-cuda	11.0.194	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.267-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.267-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.267-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.8-amzn-2	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.4.8-amzn-2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.8-amzn-2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.8-amzn-2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.36.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.36.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager,

Clasificaciones	Descripción	Acciones de reconfiguración
		NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.

Clasificaciones	Descripción	Acciones de reconfiguración
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.

Clasificaciones	Descripción	Acciones de reconfiguración
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
hudi-defaults	Cambie los valores en el archivo hudi-defaults.conf de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Livy Server and MapReduce-HistoryS erver.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Registro de cambios de la versión 5.36.1

Registro de cambios de la versión 5.36.1 y notas de la versión

Date	Evento	Descripción
2023-07-26	Actualización	Nuevas etiquetas de lanzamiento del sistema operativo `2.0.20230612.0` y `2.0.20230628.0` .
2023-05-25	Implementación completa	Amazon EMR 5.36.1 se ha implementado completamente en todas las regiones compatibles
2023-05-09	Publicación de documentos	Las notas de la versión 5.36.1 de Amazon EMR se publicaron por primera vez
2023-05-04	Versión inicial	Amazon EMR 5.36.1 se implementó por primera vez en regiones comerciales limitadas

Amazon EMR, versión 5.36.0

Versiones de las aplicaciones de la versión 5.36.0

Esta versión incluye las siguientes aplicaciones: Flink,,,, Ganglia,, HBase, HCatalog,, Hadoop, Hive, Hudi,, Hue, Iceberg,, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, MahoutOozie Phoenix Pig Presto Spark, Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
AWS SDK for Java	1.12.206	1.12.159	1.11.970	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.8	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.8	2.3.7
Hudi	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,267	0,266	0,261	0,245.1
Spark	2.4.8	2.4.8	2.4.8	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.36.0

Las siguientes notas de la versión incluyen información sobre la versión 5.36.0 de Amazon EMR. Los cambios son respecto a la versión 5.35.0.

Fecha de lanzamiento inicial: 15 de junio de 2022

Nuevas características

- La versión 5.36.0 de Amazon EMR agrega compatibilidad con el lenguaje de definición de datos (DDL) con Apache Spark en clústeres habilitados para Apache Ranger. Esto le permite utilizar Apache Ranger para administrar el acceso a operaciones como la creación, modificación y eliminación de bases de datos y tablas de un clúster de Amazon EMR.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda) , Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), EE.UU. Oeste (Oregón), EE.UU. Oeste (Norte de Californi a), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Europa (España), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (Milán), Europa (París), Europa (Estocolm o), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderaba d), Asia Pacífico (Yakarta) , Asia Pacífico (Melbourn e), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Est e) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Est e) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Est e) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Est e) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Est e) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Est e) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Est e) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos Árabes Unidos), Europa

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres) , Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos Árabes Unidos), Europa

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv)
2.0.2023 504.1	4.14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 119,1	4,14.301	3 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	22 de diciembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 912,1	4,14,291	7 de octubre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2022 719,0	4,14,287	10 de agosto de 2022	Oeste de EE. UU. (Norte de California), Europa (París), Europa (Estocolm o), Europa (Fráncfort), Asia Pacífico (Bombay), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 426,0	4,14,281	14 de junio de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de Californi a), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

Cambios, mejoras y problemas resueltos

- Las actualizaciones de Amazon EMR 5.36.0 ahora son compatibles con: aws-sdk 1.12.206, Hadoop 2.10.1-amzn-4, Hive 2.3.9-amzn-2, Hudi 0.10.1-amzn-1, Spark 2.4.8-amzn-2, Presto 0.267-amzn-1, conector de Amazon Glue 1.18.0 y EMRFS 2.51.0.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versiones de los componentes de la versión 5.36.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.16.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-notebook-env	1.5.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.21.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.7.0	Conector S3Select de EMR
emrfs	2,51,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.14.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.14.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.10.1-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.10.1-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.3.9-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.9-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.9-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.9-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.9-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.9-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.9-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.10.1-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.10.1-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.

Componente	Versión	Descripción
hudi-presto	0.10.1-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.13.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68	Servidor de bases de datos MySQL.
nvidia-cuda	11.0.194	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.

Componente	Versión	Descripción
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.267-amzn-1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.267-amzn-1	Service para ejecutar partes de una consulta.
presto-client	0.267-amzn-1	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger

Componente	Versión	Descripción
spark-client	2.4.8-amzn-2	Clientes de línea de comando de Spark.
spark-history-server	2.4.8-amzn-2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.8-amzn-2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.8-amzn-2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.36.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.36.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 5.35.0

Versiones de las aplicaciones de la versión 5.35.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, Iceberg, JupyterEnterpriseGateway, JupyterHub, Livy,, MXNet, Mahout, Oozie, PhoenixPigPresto Spark, Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
AWS SDK for Java	1.12.159	1.11.970	1.11.970	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.8	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.8	2.3.7	2.3.7
Hudi	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	-	-	-	-

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,266	0,261	0,245.1	0,245.1
Spark	2.4.8	2.4.8	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.35.0

Esta es la nota de la versión 5.35.0 de Amazon EMR.

Las siguientes notas de la versión incluyen información sobre la versión 5.35.0 de Amazon EMR. Los cambios son respecto a la versión 5.34.0.

Fecha de lanzamiento inicial: 30 de marzo de 2022

Nuevas características

- Las aplicaciones de Amazon EMR, versión 5.35, que utilizan Log4j 1.x y Log4j 2.x se actualizan para utilizar Log4j 1.2.17 (o posterior) y Log4j 2.17.1 (o posterior), respectivamente, y no requieren el uso de acciones de arranque para mitigar los problemas de CVE de las versiones anteriores. Consulte Método para mitigar CVE-2021-44228.

Cambios, mejoras y problemas resueltos

Cambios en Flink

Cambio de tipo	Descripción
Actualizaciones	• Actualice la versión de Flink a la 1.14.2. • log4j se actualizó a la versión 2.17.1.

Cambios en Hadoop

Cambio de tipo	Descripción
Portabilidad con versiones anteriores de código abierto de Hadoop desde EMR 5.34.0	• YARN-10438: Gestiona un containerID nulo en el cliente # () RMService getContainerReport • YARN-7266: los subprocesos del controlador de eventos del servidor Timeline están bloqueados • YARN-10438: ATS 1.5 no se inicia si los archivos están dañados o faltan RollingLevelDb • HADOOP-13500: sincronización de iteración del objeto de propiedades de configuración

Cambio de tipo	Descripción
	• YARN-10651: se estrelló con el NPE activado. CapacityScheduler AbstractY arnScheduler updateNodeResource() • HDFS-12221: Sustituya los xerces en XmlEditsVisitor • HDFS-16410: análisis XML inseguro en OfflineEditsXmlLoader
Cambios y correcciones de Hadoop	• Tomcat utilizado en KMS y HTTPFS se ha actualizado a la versión 8.5.75 • En la FileSystemOptimizedCommitter versión 2, el marcador de éxito se escribía en la ruta de salida de CommitJob definida al crear el confirmador. Como las rutas de salida de commitJob y en el nivel de tarea pueden diferir, se corrigió la ruta para usar la definida en los archivos de manifiesto. En el caso de los trabajos de Hive, esto hace que el marcador de éxito se escriba correctamente al realizar operaciones como la partición dinámica o UNION ALL.

Cambios en Hive

Cambio de tipo	Descripción
Hive se ha actualizado a la versión 2.3.9 de código abierto, que incluye estas correcciones de JIRA	• HIVE-17155: findConfFile () en HiveConf .java tiene algunos problemas con la ruta de configuración • HIVE-24797: deshabilite la validación de valores predeterminados al analizar los esquemas de Avro

Cambio de tipo	Descripción
	• HIVE-21563: Mejore el rendimiento de la tabla # deshabilitándola una vez getEmptyTable registerAllFunctions • HIVE-18147: Las pruebas pueden fallar con java.net. BindException: La dirección ya está en uso • HIVE-24608: cambie de nuevo a get_table en el cliente de HMS para Hive 2.3.x • HIVE-21200: Vectorización: columna de fecha que arroja java.lang. UnsupportedOperationException para parquet • HIVE-19228: elimine el uso de commons-httpclient 3.x
Backports de código abierto de Hive desde EMR 5.34.0	• HIVE-19990: la consulta con un literal de intervalo en la condición de unión falla • HIVE-25824: actualice branch-2.3 a log4j 2.17.0 • TEZ-4062: la programación de intentos especulativos debe anularse cuando la tarea se haya completado • TEZ-4108: en condiciones de carrera NullPointerException durante una ejecución especulativa • TEZ-3918: la configuración de tez.task.log.level no funciona
Actualizaciones y correcciones de Hive	• Actualice la versión de Log4j a la 2.17.1 • Actualice la versión de ORC a la 1.4.3 • Se corrigió un punto muerto debido a un hilo de penalización en ShuffleScheduler

Cambio de tipo	Descripción
Nuevas características	• Se agregó una característica para imprimir consultas de Hive en los registros de AM. Esta opción está deshabilitada de forma predeterminada. Marca/Conf.: `tez.am.emr.print.hive.query.in.log` . Estado (predeterminado): FALSO.

Cambios en Oozie

Cambio de tipo	Descripción
Backports de código abierto de Oozie desde EMR 5.34.0	• OOZIE-3652: El lanzador de Oozie debería volver a intentar enumerar los directorios cuando aparezca NoSuchFileException

Cambios en Pig

Cambio de tipo	Descripción
Actualizaciones	• log4j se actualizó a la versión 1.2.17.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre

a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.35.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.15.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.5.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.20.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.7.0	Conector S3Select de EMR
emrfs	2.49,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.14.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.14.2	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.10.1-amzn-3	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.10.1-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.3.9-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.9-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.9-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.9-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.9-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.9-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.9-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.9.0-amzn-2	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.9.0-amzn-2	Biblioteca de paquetes para ejecutar Spark con Hudi.

Componente	Versión	Descripción
hudi-presto	0.9.0-amzn-2	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.10.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.13.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68	Servidor de bases de datos MySQL.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.

Componente	Versión	Descripción
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.266-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.266-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.266-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger

Componente	Versión	Descripción
spark-client	2.4.8-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	2.4.8-amzn-1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.8-amzn-1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.8-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.35.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.35.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 5.34.0

Versiones de las aplicaciones de la versión 5.34.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet,, Mahout, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
AWS SDK for Java	1.11.970	1.11.970	1.11.970	1.11.890
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.8	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.8	2.3.7	2.3.7	2.3.7
Hudi	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,261	0,245.1	0,245.1	0,240.1
Spark	2.4.8	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.9.0	0.9.0	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.34.0

Las siguientes notas de la versión incluyen información sobre la versión 5.34.0 de Amazon EMR. Los cambios son respecto a la versión 5.33.1.

Fecha de lanzamiento inicial: 20 de enero de 2022

Fecha de lanzamiento: 21 de marzo de 2022

Nuevas características

- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.
- [Hudi] Mejoras para simplificar la configuración de Hudi. El control de simultaneidad optimista se deshabilita de forma predeterminada.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Anteriormente, el reinicio manual del administrador de recursos en un clúster multimaestro provocaba que los daemons en el clúster de Amazon EMR, como ZooKeeper, recargaran todos los nodos perdidos o retirados previamente en el archivo znode de ZooKeeper. Esto provocaba que se superaran los límites predeterminados en ciertas situaciones. Amazon EMR ahora elimina del archivo de ZooKeeper los registros de nodos perdidos o retirados que tengan más de una hora de antigüedad y se han incrementado los límites internos.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.

- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Se ha actualizado Zeppelin a la versión 0.10.0.
- Corrección de Livy: actualizado a la versión 0.7.1
- Mejora del rendimiento de Spark: los ejecutores heterogéneos se deshabilitan cuando se anulan ciertos valores de configuración de Spark en EMR 5.34.0.
- Los servidores WebHDFS y HttpFS están deshabilitados de forma predeterminada. Puede volver a habilitar WebHDFS mediante la configuración de Hadoop, `dfs.webhdfs.enabled`. El servidor HttpFS se puede iniciar mediante `sudo systemctl start hadoop-httpfs`.

Problemas conocidos

- La característica Cuadernos de Amazon EMR que se utiliza con la suplantación de usuarios de Livy no funciona porque HttpFS está deshabilitado de forma predeterminada. En este caso, el cuaderno de EMR no puede conectarse al clúster que tiene habilitada la suplantación de Livy. La solución alternativa consiste en iniciar el servidor HttpFS antes de conectar el cuaderno de EMR al clúster mediante `sudo systemctl start hadoop-httpfs`.
- Las consultas de Hue no funcionan en Amazon EMR 6.4.0 porque el servidor HttpFS de Apache Hadoop está deshabilitado de forma predeterminada. Para usar Hue en Amazon EMR 6.4.0, inicie manualmente el servidor HttpFS en el nodo principal de Amazon EMR mediante `sudo systemctl start hadoop-httpfs` o utilice un paso de Amazon EMR.
- La característica Cuadernos de Amazon EMR que se utiliza con la suplantación de usuarios de Livy no funciona porque HttpFS está deshabilitado de forma predeterminada. En este caso, el cuaderno de EMR no puede conectarse al clúster que tiene habilitada la suplantación de Livy. La solución alternativa consiste en iniciar el servidor HttpFS antes de conectar el cuaderno de EMR al clúster mediante `sudo systemctl start hadoop-httpfs`.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.34.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.14.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.4.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.18.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.7.0	Conector S3Select de EMR
emrfs	2.48,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.13.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.13.1	Administración de recursos en nodos EMR para Apache Flink. JobManager

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	2.10.1-amzn-2	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.8-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.8-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.8-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.8-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.8-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.8-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.8-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.9.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.9.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.9.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.9.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.4.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.1-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.13.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.8.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68	Servidor de bases de datos MySQL.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.1	Cliente de línea de comando de Oozie.
oozie-server	5.2.1	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.261-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.261-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.261-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.8-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	2.4.8-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.8-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.8-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.10.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.34.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.34.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.

Clasificaciones	Descripción	Acciones de reconfiguración
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 5.33.1

Versiones de las aplicaciones de la versión 5.33.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet,, Mahout, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
AWS SDK for Java	1.11.970	1.11.970	1.11.890	1.11,890
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245.1	0,245.1	0,240.1	0,240.1
Spark	2.4.7	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.9.0	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.33.1

Las siguientes notas de la versión incluyen información sobre las versiones 5.33.0 y 5.33.1 de Amazon EMR. Los cambios son respecto a la versión 5.32.0.

Fecha de lanzamiento inicial: 19 de abril de 2021

Fecha de la última actualización: 9 de agosto de 2021

Actualizaciones

- Se ha actualizado el conector de Amazon Glue a la versión 1.15.0
- Actualizado a la versión AWS SDK para Java 1.11.970
- Se ha actualizado EMRFS a la versión 2.46.0
- Se ha actualizado EMR Goodies a la versión 2.14.0
- Se ha actualizado el servidor de registros de EMR a la versión 1.9.0
- Se ha actualizado EMR S3 Dist CP a la versión 2.18.0
- Se ha actualizado EMR Secret Agent a la versión 1.8.0
- Se ha actualizado Flink a la versión 1.12.1
- Se ha actualizado Hadoop a la versión 2.10.1-amzn-1
- Se ha actualizado Hive a la versión 2.3.7-amzn-4
- Se ha actualizado Hudi a la versión 0.7.0
- Se ha actualizado Hue a la versión 4.9.0
- Se ha actualizado OpenCV a la versión 4.5.0
- Se ha actualizado Presto a la versión 0.245.1-amzn-0
- Se ha actualizado R a la versión 4.0.2

- Se ha actualizado Spark a la versión 2.4.7-amzn-1
- Actualizado a TensorFlow la versión 2.4.1
- Se ha actualizado Zeppelin a la versión 0.9.0

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.

- Configuración de un clúster para corregir los problemas de rendimiento del servidor YARN Timeline de Apache en las versiones 1 y 1.5

  Las versiones 1 y 1.5 del servidor YARN Timeline de Apache pueden provocar problemas de rendimiento con clústeres de EMR grandes y muy activos, especialmente con `yarn.resourcemanager.system-metrics-publisher.enabled=true`, que es la configuración predeterminada de Amazon EMR. Un servidor YARN Timeline v2 de código abierto resuelve el problema de rendimiento relacionado con la escalabilidad del servidor YARN Timeline.

  Otras alternativas para este problema incluyen:

  - Configuración de yarn.resourcemanager. system-metrics-publisher.enabled=false en yarn-site.xml.
  - La habilitación de la corrección para este problema al crear un clúster, tal y como se describe a continuación.

  Las siguientes versiones de Amazon EMR contienen una corrección para este problema de rendimiento del servidor YARN Timeline.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  Para habilitar la corrección en cualquiera de las versiones de Amazon EMR especificadas anteriormente, defina estas propiedades como `true` en un archivo JSON de configuración que se pasa mediante el parámetro de comando `aws emr create-cluster`: `--configurations file://./configurations.json`. También puede habilitar la corrección mediante la interfaz de usuario de la consola de reconfiguración.

  Ejemplo del contenido del archivo configurations.json:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- El tiempo de ejecución de Spark ahora es más rápido cuando se obtienen ubicaciones de particiones del metaalmacén de Hive para llevar a cabo consultas de inserción de Spark.
- Versiones de componentes actualizadas. Para obtener una lista de las versiones de los componentes, consulte Acerca de las versiones de Amazon EMR en esta guía.
- Se instaló el paquete de SDK de AWS Java en cada clúster nuevo. Se trata de un contenedor único que contiene todos los servicios SDKs y sus dependencias, en lugar de contenedores de componentes individuales. Para más información, consulte Java SDK Bundled Dependency.
- Se corrigieron los problemas de escalado administrado en versiones anteriores de Amazon EMR y se realizaron mejoras para reducir significativamente las tasas de errores de las aplicaciones.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de AWS discusión. Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Nuevas características

- Amazon EMR es compatible con Puntos de acceso de Amazon S3, una característica de Amazon S3 que le permite administrar fácilmente el acceso a los lagos de datos compartidos. Con su alias de Puntos de acceso de Amazon S3, puede simplificar el acceso a los datos a escala en Amazon EMR. Puede utilizar los puntos de acceso Amazon S3 con todas las versiones de Amazon EMR sin coste adicional en todas AWS las regiones en las que Amazon EMR esté disponible. Para más información acerca de los puntos de acceso de Amazon S3 y los alias de punto de acceso, consulte Uso de un alias de estilo bucket para el punto de acceso en la Guía del usuario de Amazon S3.
- Amazon EMR-5.33 admite los nuevos tipos de EC2 instancias de Amazon: c5a, c5ad, c6gn, c6gd, m6gd, d3, d3en, m5zn, r5b, r6gd. Consulte Tipos de instancias admitidos.

Problemas conocidos

- Reducir el límite de «máximo de AL2 archivos abiertos» en las versiones más antiguas [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que

AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
```

```
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- En el caso de los clústeres de subredes privadas de Amazon EMR 6.3.0 y 6.2.0, no puede acceder a la interfaz de usuario web de Ganglia. Aparecerá el error “Acceso denegado (403)”. Otras páginas web UIs, como Spark, Hue JupyterHub, Zeppelin, Livy y Tez, funcionan con normalidad. El acceso a la interfaz de usuario web de Ganglia en los clústeres de subredes públicas también funciona con normalidad. Para resolver este problema, reinicie el servicio httpd en el nodo principal con `sudo systemctl restart httpd`. Este problema se ha corregido en la versión 6.4.0 de Amazon EMR.
- **Important**

  Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan automáticamente esas actualizaciones. Para obtener actualizaciones del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.
- Actualmente, la GovCloud región no admite la compatibilidad con consolas para crear una configuración de seguridad que especifique la opción de integración con AWS Ranger.

La configuración de seguridad se puede llevar a cabo con la CLI. Consulte Creación de la configuración de seguridad de EMR en la Guía de administración de Amazon EMR.

- Políticas gestionadas con ámbito de aplicación: para ajustarse a las prácticas AWS recomendadas, Amazon EMR ha introducido la versión 2 de las políticas gestionadas predeterminadas con ámbito de EMR como sustitutivas de las políticas que quedarán en desuso. Consulte Políticas administradas por Amazon EMR.
- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.33.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*-

`amzn-EmrVersion`. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.14.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.2.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.18.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.6.0	Conector S3Select de EMR
emrfs	2.46.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.12.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.

Componente	Versión	Descripción
flink-jobmanager-config	1.12.1	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-1.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-1.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-1.1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.10.1-amzn-1.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.10.1-amzn-1.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-1.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-1.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-1.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-1.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-1.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-1.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.7-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.7-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.7-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.7-amzn-4	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.7-amzn-4	Hive-hbase client.
hive-metastore-server	2.3.7-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.7-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.7.0-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.7.0-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.7.0-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.9.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.2.2	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso

Componente	Versión	Descripción
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de bases de datos MySQL.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.245.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.245.1-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.245.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.7-amzn-1.1	Clientes de línea de comando de Spark.
spark-history-server	2.4.7-amzn-1.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.7-amzn-1.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.7-amzn-1.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.9.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.33.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.33.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.

Clasificaciones	Descripción	Acciones de reconfiguración
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 5.33.0

Versiones de las aplicaciones de la versión 5.33.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet,, Mahout, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
AWS SDK for Java	1.11.970	1.11.890	1.11,890	1.11,852
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,245.1	0,240.1	0,240.1	0,238,3
Spark	2.4.7	2.4.7	2.4.7	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.33.0

Versiones de los componentes de la versión 5.33.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.14.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.2.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.18.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.6.0	Conector S3Select de EMR
emrfs	2.46.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.12.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.12.1	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.10.1-amzn-1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.10.1-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.7-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.7-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.7-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.7-amzn-4	Cliente de línea de comando de Hive.
hive-hbase	2.3.7-amzn-4	Hive-hbase client.
hive-metastore-server	2.3.7-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.7-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hudi	0.7.0-amzn-1	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.7.0-amzn-1	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.7.0-amzn-1	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.9.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.2.2	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68	Servidor de bases de datos MySQL.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.5.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.245.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.245.1-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.245.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
r	4.0.2	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.7-amzn-1	Clientes de línea de comando de Spark.
spark-history-server	2.4.7-amzn-1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.7-amzn-1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.7-amzn-1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.4.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.9.0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.33.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.33.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceM anager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.pr operties de Hadoop YARN.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-session	Cambie la configuración de Flink log4j-session.properties para la sesión. Kubernetes/ Yarn	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

Clasificaciones	Descripción	Acciones de reconfiguración
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.

Clasificaciones	Descripción	Acciones de reconfiguración
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.

Clasificaciones	Descripción	Acciones de reconfiguración
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.

Clasificaciones	Descripción	Acciones de reconfiguración
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts MapReduce-HistoryServer.

Clasificaciones	Descripción	Acciones de reconfiguración
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zeppelin-site	Cambie los ajustes de configuración de zeppelin-site.xml.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 5.32.1

Versiones de las aplicaciones de la versión 5.32.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet,, Mahout, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
AWS SDK for Java	1.11.890	1.11,890	1.11,852	1.11,852
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.0	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,240.1	0,240.1	0,238,3	0,238,3
Spark	2.4.7	2.4.7	2.4.6	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.32.1

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

Cambios, mejoras y problemas resueltos

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de AWS debate. Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.32.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.1.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.17.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.6.0	Conector S3Select de EMR
emrfs	2.45,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.11.2	Administración de recursos en nodos EMR para Apache Flink. JobManager

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-0.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-0.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-0.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-0.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	2.10.1-amzn-0.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-0.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-0.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-0.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-0.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-0.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-0.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.7-amzn-3	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.7-amzn-3	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.7-amzn-3	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.7-amzn-3	Cliente de línea de comando de Hive.
hive-hbase	2.3.7-amzn-3	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.7-amzn-3	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.7-amzn-3	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.6.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.8.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68+	Servidor de bases de datos MySQL.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.240.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.240.1-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.240.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.7-amzn-0.1	Clientes de línea de comando de Spark.
spark-history-server	2.4.7-amzn-0.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.7-amzn-0.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.7-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.3.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.32.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.32.1

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.

Clasificaciones	Descripción	Acciones de reconfiguración
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 5.32.0

Versiones de las aplicaciones de la versión 5.32.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet,, Mahout, Oozie, Phoenix, PigPrestoSpark Sqoop, TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
AWS SDK for Java	1.11.890	1.11,852	1.11,852	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.6
Hadoop	2.10.1	2.10.0	2.10.0	2.8.5
Hive	2.3.7	2.3.7	2.3.7	2.3.6

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating
Hue	4.8.0	4.7.1	4.7.1	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,240.1	0,238,3	0,238,3	0.232
Spark	2.4.7	2.4.6	2.4.6	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.32.0

Las siguientes notas de la versión incluyen información sobre la versión 5.32.0 de Amazon EMR. Los cambios son respecto a la versión 5.31.0.

Fecha de lanzamiento inicial: 8 de enero de 2021

Actualizaciones

- Se ha actualizado el conector de Amazon Glue a la versión 1.14.0
- Se actualizó el SDK de Amazon SageMaker Spark a la versión 1.4.1
- Se actualizó AWS SDK para Java a la versión 1.11.890
- Se ha actualizado el conector de DynamoDB para EMR a la versión 4.16.0
- Se ha actualizado EMRFS a la versión 2.45.0
- Se han actualizado las métricas de análisis de registros de EMR a la versión 1.18.0
- MetricsAndEventsApiGateway Cliente EMR actualizado a la versión 1.5.0
- Se ha actualizado el servidor de registros de EMR a la versión 1.8.0
- Se ha actualizado EMR S3 Dist CP a la versión 2.17.0
- Se ha actualizado EMR Secret Agent a la versión 1.7.0
- Se ha actualizado Flink a la versión 1.11.2
- Se ha actualizado Hadoop a la versión 2.10.1-amzn-0
- Se ha actualizado Hive a la versión 2.3.7-amzn-3
- Se ha actualizado Hue a la versión 4.8.0
- Se ha actualizado MXNet a la versión 1.7.0
- Se ha actualizado OpenCV a la versión 4.4.0
- Se ha actualizado Presto a la versión 0.240.1-amzn-0
- Se ha actualizado Spark a la versión 2.4.7-amzn-0
- Actualizado a la versión TensorFlow 2.3.1

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Versiones de componentes actualizadas.
- Para obtener una lista de las versiones de los componentes, consulte Acerca de las versiones de Amazon EMR en esta guía.

Nuevas características

- A partir de Amazon EMR 5.32.0 y 6.5.0, el tamaño dinámico del ejecutor para Apache Spark se habilita de forma predeterminada. Para activar o desactivar esta característica, puede utilizar el parámetro de configuración `spark.yarn.heterogeneousExecutors.enabled`.

- Estado de soporte de Instance Metadata Service (IMDS) V2: los componentes de Amazon EMR 5.23.1, 5.27.1 y 5.32 o versiones posteriores se utilizan para todas las llamadas de IMDS. IMDSv2 Para las llamadas de IMDS en el código de su aplicación, puede usar ambos o configurar el IMDS para que se use solo con el IMDSv1 fin de IMDSv2 aumentar la seguridad. IMDSv2 En otras versiones de EMR 5.x, la IMDSv1 desactivación provoca un error al iniciar el clúster.
- A partir de Amazon EMR 5.32.0, puede iniciar un clúster que se integre de forma nativa con Apache Ranger. Apache Ranger es un marco de código abierto para habilitar, supervisar y administrar la seguridad integral de los datos en toda la plataforma Hadoop. Para obtener más información, consulte Apache Ranger. Con la integración nativa, puede utilizar su propio Apache Ranger para aplicar un control de acceso a los datos detallado en Amazon EMR. Consulte Integración de Amazon EMR con Apache Ranger en la Guía de versiones de Amazon EMR.
- La versión 5.32.0 de Amazon EMR es compatible con Amazon EMR en EKS. Para más información sobre cómo empezar a utilizar EMR en EKS, consulte ¿Qué es Amazon EMR en EKS?.
- La versión 5.32.0 de Amazon EMR es compatible con Amazon EMR Studio (versión preliminar). Para más información sobre cómo empezar a usar EMR Studio, consulte Amazon EMR Studio (versión preliminar).
- Políticas gestionadas con ámbito de aplicación: para ajustarse a las prácticas AWS recomendadas, Amazon EMR ha introducido la versión 2 de las políticas gestionadas predeterminadas con ámbito de EMR como sustitutivas de las políticas que quedarán en desuso. Consulte Políticas administradas por Amazon EMR.

Problemas conocidos

- En el caso de los clústeres de subredes privadas de Amazon EMR 6.3.0 y 6.2.0, no puede acceder a la interfaz de usuario web de Ganglia. Aparecerá el error “Acceso denegado (403)”. Otras páginas web UIs, como Spark, Hue, Zeppelin, Livy y JupyterHub Tez, funcionan con normalidad. El acceso a la interfaz de usuario web de Ganglia en los clústeres de subredes públicas también funciona con normalidad. Para resolver este problema, reinicie el servicio httpd en el nodo principal con `sudo systemctl restart httpd`. Este problema se ha corregido en la versión 6.4.0 de Amazon EMR.
- Reducir el límite de «máximo de archivos abiertos» en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos»

cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
```

```
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan automáticamente esas actualizaciones. Para obtener actualizaciones del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.

- Actualmente, la GovCloud región no admite la compatibilidad con consolas para crear una configuración de seguridad que especifique la opción de integración con AWS Ranger. La configuración de seguridad se puede llevar a cabo con la CLI. Consulte Creación de la configuración de seguridad de EMR en la Guía de administración de Amazon EMR.
- Cuando se AtRestEncryption habilita el cifrado HDFS en un clúster que usa Amazon EMR 5.31.0 o 5.32.0, las consultas de Hive producen la siguiente excepción de tiempo de ejecución.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
```

```
org.apache.hadoop.security.token.TokenIdentifier: Provider
org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.32.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.1	SDK de Amazon SageMaker Spark
emr-ddb	4.16.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.1.0	Entorno de Conda para cuaderno de EMR que incluye una puerta de enlace empresarial de Jupyter
emr-s3-dist-cp	2.17.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.6.0	Conector S3Select de EMR
emrfs	2.45,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.11.2	Administración de recursos en nodos EMR para Apache Flink. JobManager

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.1-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.1-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.1-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.1-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	2.10.1-amzn-0	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.1-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.1-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.1-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.1-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.1-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.1-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.7-amzn-3	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.7-amzn-3	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.7-amzn-3	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.7-amzn-3	Cliente de línea de comando de Hive.
hive-hbase	2.3.7-amzn-3	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.7-amzn-3	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.7-amzn-3	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.6.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.8.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.7.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.68	Servidor de bases de datos MySQL.
nvidia-cuda	10.1.243	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.240.1-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.240.1-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.240.1-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.7-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	2.4.7-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.7-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.7-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.3.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.32.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Las acciones de reconfiguración se producen cuando se especifica una configuración para los grupos de instancias de un clúster en ejecución. Amazon EMR solo inicia acciones de reconfiguración para las clasificaciones que se modifican. Para obtener más información, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Clasificaciones de emr-5.32.0

Clasificaciones	Descripción	Acciones de reconfiguración
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.	Restarts the ResourceManager service.
container-executor	Cambie los valores en el archivo container-executor.cfg de Hadoop YARN.	Not available.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.	Not available.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Cambie la configuración relacionada con el docker.	Not available.
emrfs-site	Cambiar la configuración de EMRFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the

Clasificaciones	Descripción	Acciones de reconfiguración
		Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	Cambiar la configuración de flink-conf.yaml.	Restarts Flink history server.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.	Restarts Flink history server.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.	Not available.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.	Restarts Flink history server.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop	Not available.
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
hbase	Configuraciones seleccion adas por Amazon EMR para Apache. HBase	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	Cambie los valores en su entorno HBase.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log 4j.properties.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hb ase.properties. HBase	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.	Not available.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.	Restarts the HBase services RegionServer, HBaseMast er, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.	Should not be reconfigured.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

Clasificaciones	Descripción	Acciones de reconfiguración
hcatalog-env	Cambie los valores en HCatalog el entorno.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.pro perties de Hive.	Not available.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.pr operties de Hive.	Not available.
hive-env	Cambiar los valores en el entorno de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.	Restarts HiveServer2 and HiveMetastore.

Clasificaciones	Descripción	Acciones de reconfiguración
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log	Not available.
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.	Not available.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.	Not available.
hue-ini	Cambiar los valores en el archivo ini de Hue	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.	Restarts Hadoop Httpfs service.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.	Not available.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.	Restarts Hadoop-KMS service.

Clasificaciones	Descripción	Acciones de reconfiguración
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.	Not available.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Cambiar los valores en el entorno de Hudi.	Not available.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.	Not available.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.	Not available.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.	Not available.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.	Not available.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.	Restarts Livy Server.
livy-env	Cambiar los valores en el entorno de Livy.	Restarts Livy Server.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.	Restarts Livy Server.

Clasificaciones	Descripción	Acciones de reconfiguración
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Cambiar los valores en el entorno de Oozie.	Restarts Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.	Restarts Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.	Restarts Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.	Not available.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.	Not available.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.	Restarts Phoenix-QueryServer.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.	Not available.
pig-env	Cambiar los valores en el entorno de Pig.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.	Restarts Oozie.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.	Not available.
presto-log	Cambiar los valores en el archivo log.properties de Presto.	Restarts Presto-Server.
presto-config	Cambiar los valores en el archivo config.properties de Presto.	Restarts Presto-Server.
presto-password-authenticator	Cambiar los valores en el archivo password-authentic ator.properties de Presto.	Not available.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.	Restarts Presto-Server.
presto-node	Cambiar valores en el archivo node.properties de Presto.	Not available.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.	Not available.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.	Not available.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.	Restarts Presto-Server.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.	Not available.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.	Not available.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.	Not available.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.	Not available.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.	Not available.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.	Not available.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.	Not available.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.	Not available.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.	Not available.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.	Not available.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.	Not available.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.	Restarts Ranger KMS Server.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.	Not available.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.	Not available.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer	Restarts EMR record server.
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.	Restarts EMR record server.

Clasificaciones	Descripción	Acciones de reconfiguración
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.	Restarts EMR record server.
spark	Configuración de Amazon EMR mantenida para Apache Spark.	This property modifies spark-defaults. See actions there.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.	Restarts Spark history server and Spark thrift server.
spark-env	Cambiar los valores en el entorno de Spark.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.	Not available.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.	Restarts Spark history server and Spark thrift server.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.	Restarts Spark history server and Spark thrift server.
sqoop-env	Cambiar los valores en el entorno de Sqoop.	Not available.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.	Not available.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.	Not available.

Clasificaciones	Descripción	Acciones de reconfiguración
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.	Restarts Oozie.
yarn-env	Cambiar los valores en el entorno de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.	Restarts Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.	Restarts Zookeeper server.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.	Restarts Zookeeper server.

Amazon EMR, versión 5.31.1

Versiones de las aplicaciones de la versión 5.31.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
AWS SDK for Java	1.11.852	1.11,852	1.11.759	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.6	2.3.6
Hadoop	2.10.0	2.10.0	2.8.5	2.8.5
Hive	2.3.7	2.3.7	2.3.6	2.3.6

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.7.1	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0,238,3	0.232	0.232
Spark	2.4.6	2.4.6	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.31.1

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

Cambios, mejoras y problemas resueltos

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket

de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de AWS debate. Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.31.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*-

`amzn-EmrVersion`. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.0	SDK de Amazon SageMaker Spark
emr-ddb	4.15.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.15.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.6.0	Conector S3Select de EMR
emrfs	2.43,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.11.0	Administración de recursos en nodos EMR para Apache Flink. JobManager

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.0-amzn-0.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.0-amzn-0.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.0-amzn-0.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.0-amzn-0.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	2.10.0-amzn-0.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.0-amzn-0.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.10.0-amzn-0.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.0-amzn-0.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.0-amzn-0.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.0-amzn-0.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.0-amzn-0.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.7-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.7-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.7-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.7-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.7-amzn-1	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.7-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.7-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.6.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.7.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.6.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.3.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.238.3-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.238.3-amzn-0	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.238.3-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.6-amzn-0.1	Clientes de línea de comando de Spark.
spark-history-server	2.4.6-amzn-0.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.6-amzn-0.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.6-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	2.1.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.31.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.31.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.

Clasificaciones	Descripción
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.

Clasificaciones	Descripción
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.

Clasificaciones	Descripción
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.pro perties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.pro perties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.31.0

Versiones de las aplicaciones de la versión 5.31.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
AWS SDK for Java	1.11.852	1.11.759	1.11.759	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.10.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.6	2.3.6	2.3.6
Hadoop	2.10.0	2.8.5	2.8.5	2.8.5
Hive	2.3.7	2.3.6	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.6.0	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0,238,3	0.232	0.232	0.232
Spark	2.4.6	2.4.5	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.31.0

Las siguientes notas de la versión incluyen información sobre la versión 5.31.0 de Amazon EMR. Los cambios son respecto a la versión 5.30.1.

Fecha de lanzamiento inicial: 9 de octubre de 2020

Fecha de la última actualización: 15 de octubre de 2020

Actualizaciones

- Se ha actualizado el conector de Amazon Glue a la versión 1.13.0
- Se actualizó el SDK de Amazon SageMaker Spark a la versión 1.4.0
- Se ha actualizado el conector Amazon Kinesis a la versión 3.5.9
- Se actualizó AWS SDK para Java a la versión 1.11.852
- Se ha actualizado Bigtop-tomcat a la versión 8.5.56
- Se ha actualizado EMRFS a la versión 2.43.0
- MetricsAndEventsApiGateway Cliente EMR actualizado a la versión 1.4.0
- Se ha actualizado EMR S3 Dist CP a la versión 2.15.0
- Se ha actualizado EMR S3 Select a la versión 1.6.0
- Se ha actualizado Flink a la versión 1.11.0
- Se ha actualizado Hadoop a la versión 2.10.0
- Se ha actualizado Hive a la versión 2.3.7
- Se ha actualizado Hudi a la versión 0.6.0
- Se ha actualizado Hue a la versión 4.7.1
- Actualizado a la versión JupyterHub 1.1.0
- Se ha actualizado MXNet a la versión 1.6.0
- Se ha actualizado OpenCV a la versión 4.3.0
- Se ha actualizado Presto a la versión 0.238.3
- Actualizado TensorFlow a la versión 2.1.0

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.

- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Las estadísticas de columnas de Hive son compatibles con las versiones 5.31.0 y posteriores de Amazon EMR.
- Versiones de componentes actualizadas.
- Compatibilidad con EMRFS S3EC V2 en Amazon EMR 5.31.0. En las versiones 1.11.837 y posteriores del SDK para Java de S3, se ha introducido la versión 2 del cliente de cifrado (S3EC V2) con varias mejoras de seguridad. Para obtener más información, consulte los siguientes temas:
  - Publicación en el blog de S3: Updates to the Amazon S3 encryption client.
  - AWS SDK para Java Guía para desarrolladores: migre los clientes de cifrado y descifrado a la versión 2.
  - Guía de administración de EMR: Cifrado del cliente de Amazon S3.

  La versión 1 del cliente de cifrado todavía está disponible en el SDK por motivos de compatibilidad con versiones anteriores.

Nuevas características

- Reducir el límite de «máximo de archivos abiertos» en los más antiguos AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

  Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

  Establecimiento explícito de un ulimit desde la línea de comandos

  1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reinicie InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Con Amazon EMR 5.31.0, puede lanzar un clúster que se integre con Lake Formation. Esta integración proporciona un filtrado de datos detallado a nivel de columna para las bases de datos y tablas del catálogo de datos de Glue. AWS También permite un inicio de sesión único federado en cuadernos de EMR o Apache Zeppelin desde un sistema de identidades empresariales. Para más información, consulte Integración de Amazon EMR con AWS Lake Formation en la Guía de administración de Amazon EMR.

  Amazon EMR with Lake Formation está disponible actualmente en 16 AWS regiones: EE.UU. Este (Ohio y Norte de Virginia), EE.UU. Oeste (Norte de California y Oregón), Asia Pacífico (Bombay, Seúl, Singapur, Sídney y Tokio), Canadá (Centro), Europa (Fráncfort, Irlanda, Londres, París y Estocolmo) y Sudamérica (São Paulo).

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las

operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Cuando se AtRestEncryption habilita el cifrado HDFS en un clúster que usa Amazon EMR 5.31.0 o 5.32.0, las consultas de Hive producen la siguiente excepción de tiempo de ejecución.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:

- Se analizan dos o más particiones de la misma tabla.
- Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
- El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.31.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.4.0	SDK de Amazon SageMaker Spark

Componente	Versión	Descripción
emr-ddb	4.15.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.15.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.6.0	Conector S3Select de EMR
emrfs	2.43,0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.11.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
flink-jobmanager-config	1.11.0	Administración de recursos en nodos EMR para Apache Flink. JobManager
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.10.0-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.10.0-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.10.0-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.10.0-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.10.0-amzn-0	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.10.0-amzn-0	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.10.0-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.10.0-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.10.0-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.10.0-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.10.0-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.7-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.7-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.7-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.7-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.7-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.7-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.7-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hudi	0.6.0-amzn-0	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-spark	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Spark con Hudi.
hudi-presto	0.6.0-amzn-0	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.7.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.6.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	4.3.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.238.3-amzn-0	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.238.3-amzn-0	Service para ejecutar partes de una consulta.
presto-client	0.238.3-amzn-0	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.6-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	2.4.6-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.6-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.6-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	2.1.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.31.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.31.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.

Clasificaciones	Descripción
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.

Clasificaciones	Descripción
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.

Clasificaciones	Descripción
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer

Clasificaciones	Descripción
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.30.2

Versiones de las aplicaciones de la versión 5.30.2

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
AWS SDK for Java	1.11.759	1.11.759	1.11.759	1.11.682
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.6

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.10.0	1.9.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.6.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.232	0.227
Spark	2.4.5	2.4.5	2.4.5	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.30.2

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

Cambios, mejoras y problemas resueltos

- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.

- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de AWS debate. Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre

a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.30.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.3.0	SDK de Amazon SageMaker Spark
emr-ddb	4.14.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.14.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.5.0	Conector S3Select de EMR
emrfs	2.40.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.10.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-6.1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-6.1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-6.1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6.1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-6.1	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-6.1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-6.1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-6.1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.5-amzn-6.1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-6.1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-6.1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.6-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.6-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.6-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.6-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.6-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.6-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.6-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.2-incubating	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.2-incubating	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.6.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.232	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.232	Service para ejecutar partes de una consulta.
presto-client	0.232	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.5-amzn-0.1	Clientes de línea de comando de Spark.
spark-history-server	2.4.5-amzn-0.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.

Componente	Versión	Descripción
spark-on-yarn	2.4.5-amzn-0.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.5-amzn-0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.30.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.30.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.

Clasificaciones	Descripción
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.

Clasificaciones	Descripción
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift. properties de Presto.

Clasificaciones	Descripción
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.

Clasificaciones	Descripción
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.30.1

Versiones de las aplicaciones de la versión 5.30.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
AWS SDK for Java	1.11.759	1.11.759	1.11.682	1.11.659
Python	2.7, 3.7	2.7, 3.7	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.9.1	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
HBase	1.4.13	1.4.13	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.227	0.227
Spark	2.4.5	2.4.5	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.30.1

Las siguientes notas de la versión incluyen información sobre la versión 5.30.1 de Amazon EMR. Los cambios son respecto a la versión 5.30.0.

Fecha de lanzamiento inicial: 30 de junio de 2020

Fecha de la última actualización: 24 de agosto de 2020

Cambios, mejoras y problemas resueltos

- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Se ha solucionado el problema que provocaba que el proceso del controlador de instancias generaba un número infinito de procesos.
- Se ha solucionado el problema que provocaba que Hue no pudiera ejecutar una consulta de Hive, mostrara el mensaje “database is locked” e impidiera la ejecución de consultas.
- Se ha corregido un problema de Spark que permitía ejecutar más tareas simultáneamente en el clúster de EMR.
- Se ha corregido un error en el cuaderno de Jupyter que provocaba el “too many files open” en el servidor de Jupyter.
- Se ha corregido un problema con las horas de inicio de los clústeres.

Nuevas características

- Las interfaces de aplicación persistentes de la IU de Tez y el servidor YARN Timeline están disponibles con las versiones 6.x de Amazon EMR y la versión 5.30.1 y posteriores de EMR. El acceso mediante un enlace al historial de aplicaciones persistente con un solo clic le permite acceder rápidamente al historial de trabajos sin necesidad de configurar un proxy web a través de una conexión SSH. Los registros de los clústeres activos y terminados están disponibles durante los 30 días posteriores a la finalización de la aplicación. Para más información, consulte Ver interfaces de usuario de aplicaciones persistentes en la Guía de administración de Amazon EMR.
- La ejecución de cuadernos EMR APIs está disponible para ejecutar cuadernos EMR mediante un script o una línea de comandos. La capacidad de iniciar, detener, enumerar y describir las ejecuciones de cuadernos EMR sin la AWS consola le permite controlar mediante programación un cuaderno EMR. Al utilizar una celda de cuaderno parametrizada, puede pasar diferentes valores de parámetros a un cuaderno sin tener que crear una copia del cuaderno para cada nuevo conjunto de valores de parámetros. Consulte Acciones de la API de EMR. Para ver un código de ejemplo, consulte Ejemplos de comandos para ejecutar Cuadernos de EMR mediante programación.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
```

```
sudo systemctl daemon-reload
```

- EMR Notebooks

  La característica que le permite instalar kernels y bibliotecas de Python adicionales en el nodo principal del clúster está deshabilitada de forma predeterminada en la versión 5.30.1 de EMR. Para más información sobre esta característica, consulte Instalación de kernels y bibliotecas de Python en el nodo principal del clúster.

  Para habilitar la característica, haga lo siguiente:

  1. Asegúrese de que la política de permisos adjunta al rol de servicio de Cuadernos de EMR permita la siguiente acción:

     ```
     elasticmapreduce:ListSteps
     ```

     Para obtener más información, consulte Rol de servicio para cuadernos de EMR.

  2. Utilice el AWS CLI para ejecutar un paso en el clúster que configura los EMR Notebooks, como se muestra en el siguiente ejemplo. *us-east-1*Sustitúyalo por la región en la que reside el clúster. Para obtener más información, consulte Adición de pasos a un clúster mediante la AWS CLI.

     ```
     aws emr add-steps  --cluster-id MyClusterID --steps
      Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
     east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
     awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
     ```

- Escalado administrado

  Las operaciones de escalado administrado en los clústeres 5.30.0 y 5.30.1 sin Presto instalado pueden provocar errores en las aplicaciones o provocar que un grupo de instancias o una flota de instancias uniformes permanezcan en estado ARRESTED, especialmente cuando una operación de reducción vertical va seguida inmediatamente de una operación de escalado vertical.

  Como solución alternativa, elija Presto como aplicación para instalar cuando cree un clúster con las versiones 5.30.0 y 5.30.1 de Amazon EMR, incluso si su trabajo no requiere Presto.

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las

operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros

14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.30.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.3.0	SDK de Amazon SageMaker Spark
emr-ddb	4.14.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.14.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.5.0	Conector S3Select de EMR
emrfs	2.40.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.10.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-6	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-6	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.5-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.6-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.6-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.6-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.6-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.6-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.6-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.6-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.2-incubating	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.2-incubating	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.6.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.232	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.232	Service para ejecutar partes de una consulta.
presto-client	0.232	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.5-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	2.4.5-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.

Componente	Versión	Descripción
spark-on-yarn	2.4.5-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.5-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.30.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.30.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.

Clasificaciones	Descripción
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.

Clasificaciones	Descripción
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.

Clasificaciones	Descripción
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.

Clasificaciones	Descripción
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.30.0

Versiones de las aplicaciones de la versión 5.30.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
AWS SDK for Java	1.11.759	1.11.682	1.11.659	1.11.659
Python	2.7, 3.7	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.9.1	1.9.0	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
HBase	1.4.13	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.0.0	1.0.0	1.0.0
Livy	0.7.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.227	0.227	0.227
Spark	2.4.5	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.30.0

Las siguientes notas de la versión incluyen información sobre la versión 5.30.0 de Amazon EMR. Los cambios son respecto a la versión 5.29.0.

Fecha de lanzamiento inicial: 13 de mayo de 2020

Fecha de la última actualización: 25 de junio de 2020

Actualizaciones

- Actualizado a la versión AWS SDK para Java 1.11.759
- Se actualizó el SDK de Amazon SageMaker Spark a la versión 1.3.0
- Se actualizó el servidor de registros de EMR a la versión 1.6.0
- Se ha actualizado Flink a la versión 1.10.0
- Se actualizó Ganglia a la versión 3.7.2
- Se actualizó HBase a la versión 1.4.13
- Se actualizó Hudi a la versión 0.5.2-incubating
- Se actualizó Hue a la versión 4.6.0
- Actualizado a la versión JupyterHub 1.1.0
- Se actualizó Livy a la versión 0.7.0-incubating
- Se actualizó Oozie a la versión 5.2.0
- Se actualizó Presto a la versión 0.232
- Se actualizó Spark a la versión 2.4.5
- Conectores y controladores actualizados: Amazon Glue Connector 1.12.0; Amazon Kinesis Connector 3.5.0; EMR DynamoDB Connector 4.14.0

Nuevas características

- Cuadernos de EMR: cuando se utiliza con clústeres de EMR creados con la versión 5.30.0, los kernels de cuadernos de EMR se ejecutan en un clúster. Esto mejora el rendimiento del bloc de notas y le permite instalar y personalizar kernels. También puede instalar bibliotecas de Python en el nodo principal del clúster. Para obtener más información, consulte Instalación y uso de kernels y bibliotecas en la Guía de administración de EMR.
- Escalado administrado: con la versión 5.30.0 y posteriores de Amazon EMR, puede habilitar el escalado administrado de EMR para aumentar o disminuir automáticamente el número de instancias o unidades del clúster en función de la carga de trabajo. Amazon EMR evalúa continuamente las métricas del clúster para tomar decisiones de escalado que optimicen los clústeres en cuanto al costo y la velocidad. Para más información, consulte Recursos de escalado de clúster en la Guía de administración de Amazon EMR.
- Cifrar los archivos de registro almacenados en Amazon S3: con Amazon EMR versión 5.30.0 y versiones posteriores, puede cifrar los archivos de registro almacenados en Amazon S3 con una clave gestionada por el cliente. AWS KMS Para más información, consulte Consultar archivos de registros almacenados en Amazon S3 en la Guía de administración de Amazon EMR.
- Compatibilidad con Amazon Linux 2: en la versión 5.30.0 y posteriores de EMR, EMR utiliza el sistema operativo Amazon Linux 2. La nueva personalización AMIs (Amazon Machine Image) debe estar basada en la AMI de Amazon Linux 2. Para obtener más información, consulte Uso de una AMI personalizada.
- Escalado automático estable de Presto: los clústeres de EMR que utilizan la versión 5.30.0 se pueden configurar con un periodo de espera de escalado automático que da tiempo a las tareas de Presto para terminar de ejecutarse antes de que su nodo se retire. Para obtener más información, consulte Uso del escalado automático de Presto con la retirada estable.
- Creación de instancias de flota con una nueva opción de estrategia de asignación: hay una nueva opción de estrategia de asignación disponible en la versión 5.12.1 y posteriores de EMR. Ofrece un aprovisionamiento de clústeres más rápido, una asignación de spot más precisa y menos interrupciones de las instancias de spot. Se requieren actualizaciones de los roles de servicio de EMR no predeterminados. Consulte Configurar flotas de instancias.
- Comandos sudo systemctl stop y sudo systemctl start: en la versión 5.30.0 y posteriores de EMR, que utilizan el sistema operativo Amazon Linux 2, EMR utiliza los comandos `sudo systemctl stop` y `sudo systemctl start` para reiniciar los servicios. Para más información, consulte ¿Cómo reinicio un servicio en Amazon EMR?.

Cambios, mejoras y problemas resueltos

- La versión 5.30.0 de EMR no instala Ganglia de forma predeterminada. Puede seleccionar explícitamente Ganglia para instalar cuando cree un clúster.
- Optimizaciones de rendimiento de Spark.
- Optimizaciones de rendimiento de Presto.
- Python 3 es la opción predeterminada para la versión 5.30.0 y posteriores de Amazon EMR.
- El grupo de seguridad administrado predeterminado para el acceso al servicio en subredes privadas se ha actualizado con nuevas reglas. Si utiliza un grupo de seguridad personalizado para el acceso al servicio, debe incluir las mismas reglas que el grupo de seguridad administrado predeterminado. Para más información, consulte Grupo de seguridad administrado por Amazon EMR para el acceso al servicio (Subredes privadas). Si utiliza un rol de servicio personalizado para Amazon EMR, debe conceder permiso a `ec2:describeSecurityGroups` para que EMR pueda validar si los grupos de seguridad se crean correctamente. Si utiliza `EMR_DefaultRole`, este permiso ya está incluido en la política administrada predeterminada.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
```

```
sudo systemctl daemon-reload
```

- Escalado administrado

  Las operaciones de escalado administrado en los clústeres 5.30.0 y 5.30.1 sin Presto instalado pueden provocar errores en las aplicaciones o provocar que un grupo de instancias o una flota de instancias uniformes permanezcan en estado `ARRESTED`, especialmente cuando una operación de reducción vertical va seguida inmediatamente de una operación de escalado vertical.

  Como solución alternativa, elija Presto como aplicación para instalar cuando cree un clúster con las versiones 5.30.0 y 5.30.1 de Amazon EMR, incluso si su trabajo no requiere Presto.

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede

configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- El motor de base de datos predeterminado para Hue 4.6.0 es SQLite, lo que provoca problemas al intentar utilizar Hue con una base de datos externa. Para solucionar este problema, defina `engine` en su clasificación de configuración `hue-ini` como `mysql`. Este problema se ha corregido en la versión 5.30.1 de Amazon EMR.
- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versiones de los componentes de la versión 5.30.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.3.0	SDK de Amazon SageMaker Spark
emr-ddb	4.14.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.13.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.5.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-notebook-env	1.0.0	Entorno Conda para emr notebook
emr-s3-dist-cp	2.14.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.5.0	Conector S3Select de EMR
emrfs	2.40.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.10.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	2.8.5-amzn-6	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-6	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.13	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.13	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.13	HBase cliente de línea de comandos.
hbase-rest-server	1.4.13	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.13	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.6-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.6-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.6-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.6-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.6-amzn-2	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.6-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.6-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.2-incubating	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.2-incubating	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.6.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.1.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.7.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mariadb-server	5.5.64	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.2.0	Cliente de línea de comando de Oozie.
oozie-server	5.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.232	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.232	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
presto-client	0.232	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.3	Proyecto R para análisis estadístico
ranger-kms-server	1.2.0	Sistema de administración de claves Apache Ranger
spark-client	2.4.5-amzn-0	Clientes de línea de comando de Spark.
spark-history-server	2.4.5-amzn-0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.5-amzn-0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.5-amzn-0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.30.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.30.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.

Clasificaciones	Descripción
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
hudi-env	Cambiar los valores en el entorno de Hudi.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.

Clasificaciones	Descripción
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.

Clasificaciones	Descripción
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.pro perties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.pro perties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.29.0

Versiones de las aplicaciones de la versión 5.29.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
AWS SDK for Java	1.11.682	1.11.659	1.11.659	1.11.615
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.1	1.9.0	1.9.0	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.227	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.29.0

Las siguientes notas de la versión incluyen información sobre la versión 5.29.0 de Amazon EMR. Los cambios son respecto a la versión 5.28.1.

Fecha de lanzamiento inicial: 17 de enero de 2020

Actualizaciones

- Actualizado a la versión AWS SDK para Java 1.11.682
- Actualización de Hive a la versión 2.3.6
- Se ha actualizado Flink a la versión 1.9.1
- Se ha actualizado EMRFS a la versión 2.38.0
- Se ha actualizado el conector EMR DynamoDB a la versión 4.13.0

Cambios, mejoras y problemas resueltos

- Spark
  - Optimizaciones de rendimiento de Spark.
- EMRFS
  - La guía de administración actualiza la configuración predeterminada de emrfs-site.xml para una visualización coherente.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.29.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.6	SDK de Amazon SageMaker Spark
emr-ddb	4.13.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.12.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.13.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.4.0	Conector S3Select de EMR
emrfs	2.38.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.9.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-5	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.5-amzn-5	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.10	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.10	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.10	HBase cliente de línea de comandos.
hbase-rest-server	1.4.10	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.10	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.6-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.6-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.6-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.6-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.6-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.6-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.6-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.0-incubating	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.0-incubating	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
jupyterhub	1.0.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.227	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.227	Service para ejecutar partes de una consulta.
presto-client	0.227	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.4	Clientes de línea de comando de Spark.
spark-history-server	2.4.4	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.4	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.4.4	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.29.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.29.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase

Clasificaciones	Descripción
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.

Clasificaciones	Descripción
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.

Clasificaciones	Descripción
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.pro perties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.pro perties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.

Clasificaciones	Descripción
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.28.1

Versiones de las aplicaciones de la versión 5.28.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
AWS SDK for Java	1.11.659	1.11.659	1.11.615	1.11.615
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.9.0	1.8.1	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.5	2.3.5

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
Hudi	0.5.0-incubating	0.5.0-incubating	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.224	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.28.1

Las siguientes notas de la versión incluyen información sobre la versión 5.28.1 de Amazon EMR. Los cambios son respecto a la versión 5.28.0.

Fecha de lanzamiento inicial: 10 de enero de 2020

Cambios, mejoras y problemas resueltos

- Spark
  - Se han corregido problemas de compatibilidad con Spark.
- CloudWatch Métricas
  - Se corrigió la publicación de Amazon CloudWatch Metrics en un clúster de EMR con varios nodos principales.
- Mensaje de registro deshabilitado
  - Se ha deshabilitado el mensaje de registro falso: "... usando una versión anterior (<4.5.8) del cliente http de Apache".

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.28.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.6	SDK de Amazon SageMaker Spark
emr-ddb	4.12.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.11.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.13.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR
emrfs	2.37.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.9.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-5	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.5-amzn-5	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.10	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.10	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.10	HBase cliente de línea de comandos.
hbase-rest-server	1.4.10	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.10	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.6-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.6-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.6-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.6-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.6-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.6-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.6-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.0-incubating	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.0-incubating	Biblioteca de paquetes para ejecutar Presto con Hudi.
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
jupyterhub	1.0.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.227	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.227	Service para ejecutar partes de una consulta.
presto-client	0.227	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.4	Clientes de línea de comando de Spark.
spark-history-server	2.4.4	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.4	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.4.4	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.28.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.28.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase

Clasificaciones	Descripción
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.

Clasificaciones	Descripción
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.

Clasificaciones	Descripción
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.pro perties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.pro perties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.

Clasificaciones	Descripción
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.28.0

Versiones de las aplicaciones de la versión 5.28.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi,, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, PrestoSparkSqoop TensorFlow, Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
AWS SDK for Java	1.11.659	1.11.615	1.11.615	1.11.595
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.8.1	1.8.1	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.5	2.3.5	2.3.5

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
Hudi	0.5.0-incubating	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.2	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.224	0.224	0.220
Spark	2.4.4	2.4.4	2.4.4	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.28.0

Las siguientes notas de la versión incluyen información sobre la versión 5.28.0 de Amazon EMR. Los cambios son respecto a la versión 5.27.0.

Fecha de lanzamiento inicial: 12 de noviembre de 2019

Actualizaciones

- Actualización de Flink a la versión 1.9.0
- Actualización de Hive a la versión 2.3.6
- Actualizado a la versión MXNet 1.5.1
- Actualización de Phoenix a la versión 4.14.3
- Actualización de Presto a la versión 0.227
- Actualización de Zeppelin a la versión 0.8.2

Nuevas características

- Apache Hudi ahora está disponible para que Amazon EMR lo instale al crear un clúster. Para obtener más información, consulte Hudi.
- (25 de noviembre de 2019) Ahora puede ejecutar varios pasos en paralelo para mejorar la utilización del clúster y ahorrar costos. También puede cancelar los pasos pendientes y los pasos que estén en ejecución. Para obtener más información, consulte Cómo trabajar con Steps utilizando la consola AWS CLI and.
- (3 de diciembre de 2019) Ahora puede crear y ejecutar clústeres de EMR en ellos. AWS Outposts AWS Outposts habilita AWS los servicios, la infraestructura y los modelos operativos nativos en las instalaciones locales. En AWS Outposts los entornos, puede usar las mismas AWS APIs herramientas e infraestructura que usa en la AWS nube. Para obtener más información, consulte Clústeres de EMR en. AWS Outposts
- (11 de marzo de 2020) A partir de la versión 5.28.0 de Amazon EMR, puede crear y ejecutar clústeres de Amazon EMR en una subred de Zonas AWS locales como una extensión lógica de una región que admite Zonas locales. AWS Una zona local permite que las funciones de Amazon EMR y un subconjunto de AWS servicios, como los servicios de procesamiento y almacenamiento, se ubiquen más cerca de los usuarios, lo que proporciona un acceso de muy baja latencia a las aplicaciones que se ejecutan localmente. Para obtener una lista de las zonas locales disponibles,

consulte Zonas locales de AWS. Para obtener información sobre el acceso a las zonas AWS locales disponibles, consulte Regiones, Zonas de disponibilidad y Zonas locales.

Actualmente, Zonas locales no es compatible con Cuadernos de Amazon EMR y no admite conexiones directas a Amazon EMR mediante el punto de conexión de VPC de la interfaz (AWS PrivateLink).

Cambios, mejoras y problemas resueltos

- Compatibilidad mejorada de aplicaciones para clústeres de alta disponibilidad
  - Para más información, consulte Aplicaciones admitidas en un clúster de EMR con varios nodos principales en la Guía de administración de Amazon EMR.
- Spark
  - Optimizaciones de rendimiento
- Hive
  - Optimizaciones de rendimiento
- Presto
  - Optimizaciones de rendimiento

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.28.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.6	SDK de Amazon SageMaker Spark

Componente	Versión	Descripción
emr-ddb	4.12.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.11.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.13.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR
emrfs	2.37.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.9.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-5	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.8.5-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-5	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.10	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.10	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.10	HBase cliente de línea de comandos.
hbase-rest-server	1.4.10	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.10	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.3.6-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.6-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.6-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.6-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.6-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.6-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.6-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hudi	0.5.0-incubating	Marco de procesamiento incremental para impulsar la canalización de datos a baja latencia y alta eficiencia.
hudi-presto	0.5.0-incubating	Biblioteca de paquetes para ejecutar Presto con Hudi.

Componente	Versión	Descripción
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.0.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.5.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	4.14.3- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.3- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.227	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.227	Service para ejecutar partes de una consulta.
presto-client	0.227	Cliente de línea de comandos Presto que se instala en los nodos principales en espera de un clúster de HA donde el servidor Presto no se ha iniciado.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.4	Clientes de línea de comando de Spark.
spark-history-server	2.4.4	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.

Componente	Versión	Descripción
spark-on-yarn	2.4.4	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.4	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.28.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.28.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.

Clasificaciones	Descripción
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.

Clasificaciones	Descripción
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.

Clasificaciones	Descripción
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.

Clasificaciones	Descripción
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.27.1

Versiones de las aplicaciones de la versión 5.27.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, JupyterHub, Livy, MXNet, Mahout, Oozie,, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
AWS SDK for Java	1.11.615	1.11.615	1.11.595	1.11.566
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.1	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
HBase	1.4.10	1.4.10	1.4.10	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.5
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.2	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.224	0.220	0.220
Spark	2.4.4	2.4.4	2.4.3	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.13.1	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

Notas de la versión 5.27.1

Esta es una versión de revisión. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

Estado de soporte de Instance Metadata Service (IMDS) V2: los componentes de Amazon EMR 5.23.1, 5.27.1 y 5.32 o versiones posteriores se utilizan para todas las llamadas de IMDS. IMDSv2 Para las llamadas de IMDS en el código de su aplicación, puede usar ambos o configurar el IMDS para que se use solo con el IMDSv1 fin de IMDSv2 aumentar la seguridad. IMDSv2 En otras versiones de EMR 5.x, la IMDSv1 desactivación provoca un error al iniciar el clúster.

Versiones de los componentes de la versión 5.27.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.4	SDK de Amazon SageMaker Spark
emr-ddb	4.12.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.11.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.13.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR
emrfs	2.36.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.8.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.10	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.10	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.10	HBase cliente de línea de comandos.
hbase-rest-server	1.4.10	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.10	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.3.5-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.5-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.5-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.5-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.5-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.5-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.5-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.0.0	Servidor multiusuario para blocs de notas Jupyter

Componente	Versión	Descripción
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.4.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.2- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.14.2- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.224	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.224	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.4	Clientes de línea de comando de Spark.
spark-history-server	2.4.4	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.4	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.4	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.27.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.27.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.

Clasificaciones	Descripción
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.

Clasificaciones	Descripción
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.

Clasificaciones	Descripción
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.pro perties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.pro perties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.27.0

Versiones de las aplicaciones de la versión 5.27.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, JupyterHub, Livy, MXNet, Mahout, Oozie,, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
AWS SDK for Java	1.11.615	1.11.595	1.11.566	1.11,546
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
JupyterHub	1.0.0	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.220	0.220	0.219
Spark	2.4.4	2.4.3	2.4.3	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.13.1	1.13.1	1.12.0
Tez	0.9.2	0.9.2	0.9.2	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.13

Notas de la versión 5.27.0

Las siguientes notas de la versión incluyen información sobre la versión 5.27.0 de Amazon EMR. Los cambios son respecto a la versión 5.26.0.

Fecha de lanzamiento inicial: 23 de septiembre de 2019

Actualizaciones

- AWS SDK para Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- TensorFlow 1.14.0
- Conectores y controladores:
  - Conector de DynamoDB 4.12.0

Nuevas características

- (24 de octubre de 2019) Las siguientes características nuevas de Cuadernos de EMR están disponibles en todas las versiones de Amazon EMR.
  - Puede asociar repositorios de Git con EMR Notebooks para guardar los cuadernos en un entorno con control de versiones. Puede compartir código con sus compañeros y reutilizar los cuadernos de Jupyter existentes a través de repositorios de Git remotos. Para más información, consulte Agregar repositorios de Git con Cuadernos de Amazon EMR en la Guía de administración de Amazon EMR.
  - La utilidad nbdime ahora está disponible en Cuadernos de EMR para simplificar la comparación y la combinación de cuadernos.
  - Ahora son compatibles los cuadernos EMR. JupyterLab JupyterLab es un entorno de desarrollo interactivo basado en la web totalmente compatible con los portátiles Jupyter. Ahora puede elegir abrir su bloc de notas en el editor de cuadernos de JupyterLab Jupyter.
- (30 de octubre de 2019) Con las versiones 5.25.0 y posteriores de Amazon EMR, puede conectarse a la interfaz de usuario del servidor del historial de Spark desde la página Resumen del clúster o desde la pestaña Historial de aplicaciones de la consola. En lugar de configurar un proxy web a través de una conexión SSH, puede acceder rápidamente a la interfaz de usuario del servidor de historial de Spark para ver las métricas de las aplicaciones y acceder a los archivos de registro relevantes de los clústeres activos y terminados. Para más información, consulte Acceso fuera del clúster a las interfaces de usuario de aplicaciones persistentes en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Clúster de Amazon EMR con varios nodos principales
  - Puede instalar y ejecutar Flink en un clúster de Amazon EMR con varios nodos principales. Para más información, consulte Aplicaciones y características compatibles.
  - Puede configurar el cifrado transparente de HDFS en un clúster de Amazon EMR con varios nodos principales. Para más información, consulte Cifrado transparente de HDFS en clústeres de EMR con varios nodos principales.
  - Ahora puede modificar la configuración de las aplicaciones que se ejecutan en un clúster de Amazon EMR con varios nodos principales. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.
- Conector de DynamoDB para Amazon EMR
  - El conector de DynamoDB para Amazon EMR ahora admite los siguientes tipos de datos de DynamoDB: boolean, list, map, item y null. Para más información, consulte Configurar una tabla de Hive para ejecutar comandos de Hive.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.27.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.4	SDK de Amazon SageMaker Spark
emr-ddb	4.12.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.11.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.13.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR
emrfs	2.36.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.8.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.5-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.10	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.10	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.10	HBase cliente de línea de comandos.
hbase-rest-server	1.4.10	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.10	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.5-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.5-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.5-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.5-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.5-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.5-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.5-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	1.0.0	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.4.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.2- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.2- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.224	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.224	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.4	Clientes de línea de comando de Spark.
spark-history-server	2.4.4	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.4	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.4	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.14.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.27.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.27.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.

Clasificaciones	Descripción
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log

Clasificaciones	Descripción
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.

Clasificaciones	Descripción
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.

Clasificaciones	Descripción
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores en el ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.
ranger-kms-log4j	Cambie los valores del archivo kms-log4j.properties de Ranger KMS.

Clasificaciones	Descripción
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.26.0

Versiones de las aplicaciones de la versión 5.26.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, JupyterHub, Livy, MXNet, Mahout, Oozie,, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
AWS SDK for Java	1.11.595	1.11.566	1.11,546	1.11,546
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.220	0.219	0.219
Spark	2.4.3	2.4.3	2.4.2	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.13.1	1.12.0	1.12.0
Tez	0.9.2	0.9.2	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.13	3.4.13

Notas de la versión 5.26.0

Las siguientes notas de la versión incluyen información sobre la versión 5.26.0 de Amazon EMR. Los cambios son respecto a la versión 5.25.0.

Fecha de lanzamiento inicial: 8 de agosto de 2019

Fecha de la última actualización: 19 de agosto de 2019

Actualizaciones

- AWS SDK para Java 1.11.595
- HBase 1.4,10
- Phoenix 4.14.2

- Conectores y controladores:
  - Conector de DynamoDB 4.11.0
  - MariaDB Connector 2.4.2
  - Controlador JDBC de Amazon Redshift, versión 1.2.32.1056

Nuevas características

- (Beta) Con Amazon EMR 5.26.0, puede lanzar un clúster que se integre con Lake Formation. Esta integración proporciona un acceso detallado a nivel de columnas a las bases de datos y tablas del catálogo de datos de Glue. AWS También permite un inicio de sesión único federado en cuadernos de EMR o Apache Zeppelin desde un sistema de identidades empresariales. Para obtener más información, consulte Integración de Amazon EMR con AWS Lake Formation (beta).
- (19 de agosto de 2019) El bloqueo de acceso público de Amazon EMR ya está disponible en todas las versiones de Amazon EMR que admiten grupos de seguridad. Bloquear el acceso público es una configuración que abarca toda la cuenta y que se aplica a cada región. AWS Bloquear el acceso público impide que se inicie un clúster cuando cualquier grupo de seguridad asociado al clúster tiene una regla que permite el tráfico entrante desde IPv4 0.0.0.0/0 o IPv6 : :/0 (acceso público) en un puerto, a menos que se especifique un puerto como excepción. El puerto 22 es una excepción de forma predeterminada. Para más información, consulte Uso del bloqueo de acceso público de Amazon EMR en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- EMR Notebooks
  - Con la versión 5.26.0 y posteriores de EMR, Cuadernos de EMR admite bibliotecas de Python limitadas a cuadernos, además de las bibliotecas de Python predeterminadas. Puede instalar bibliotecas limitadas a cuadernos desde el editor de cuadernos sin tener que volver a crear un clúster ni volver a asociar un cuaderno a un clúster. Las bibliotecas con ámbito de cuaderno se instalan en un entorno virtual de Python, por lo que solo se aplican a la sesión de cuaderno actual. Esto le permite aislar las dependencias de los cuadernos. Para más información, consulte Uso de bibliotecas adaptadas a los cuadernos en la Guía de administración de Amazon EMR.
- EMRFS
  - Puede habilitar una función de ETag verificación (beta) configurándola en. `fs.s3.consistent.metadata.etag.verification.enabled` `true` Con esta función,

EMRFS utiliza Amazon S3 ETags para comprobar que los objetos que se están leyendo son de la última versión disponible. Esta función resulta útil para los casos de read-after-update uso en los que los archivos de Amazon S3 se sobrescriben y conservan el mismo nombre. Esta capacidad de ETag verificación no funciona actualmente con S3 Select. Para más información, consulte Configuración de la vista consistente.

- Spark
  - Las siguientes optimizaciones ahora se habilitan de forma predeterminada: la eliminación dinámica de particiones, DISTINCT antes de INTERSECT, las mejoras en la inferencia estadística del plan de SQL para JOIN seguida de consultas DISTINCT, el aplanamiento de las subconsultas escalares, el reordenamiento optimizado de las uniones y la unión mediante filtros de Bloom. Para más información, consulte Optimización del rendimiento de Spark.
  - Se ha mejorado la generación de código en todas las etapas para Sort Merge Join.
  - Se ha mejorado la reutilización de fragmentos de consultas y subconsultas.
  - Mejoras en la preasignación de los ejecutores en el inicio de Spark.
  - Las uniones con filtro de Bloom ya no se aplican cuando el lado más pequeño de la unión incluye una sugerencia de transmisión.
- Tez
  - Se ha resuelto un problema con Tez. La interfaz de usuario de Tez ahora funciona en un clúster de Amazon EMR con varios nodos principales.

Problemas conocidos

- Las capacidades mejoradas de generación de código en todas las etapas de Sort Merge Join pueden aumentar la presión de la memoria cuando están habilitadas. Esta optimización mejora el rendimiento, pero puede provocar reintentos de trabajo o errores si `spark.yarn.executor.memory0verheadFactor` no se ajusta para proporcionar suficiente memoria. Para deshabilitar esta característica, establezca `spark.sql.sortMergeJoinExec.extendedCodegen.enabled` como “Falso”.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical

afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.26.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres

veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.4	SDK de Amazon SageMaker Spark
emr-ddb	4.11.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.10.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.12.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR
emrfs	2.35.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.8.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-4	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.8.5-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.10	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.10	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.10	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.10	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.10	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.5-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.5-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.5-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.5-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.5-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.5-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.5-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.6	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.4.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	4.14.2- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.2- -1.4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.220	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.220	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.3	Clientes de línea de comando de Spark.
spark-history-server	2.4.3	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.3	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.

Componente	Versión	Descripción
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.13.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.26.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.26.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase

Clasificaciones	Descripción
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.

Clasificaciones	Descripción
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.

Clasificaciones	Descripción
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.

Clasificaciones	Descripción
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.25.0

Versiones de las aplicaciones de la versión 5.25.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
AWS SDK for Java	1.11.566	1.11,546	1.11,546	1.11,519
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.4
Livy	0.6.0	0.6.0	0.6.0	0.5.0
MXNet	1.4.0	1.4.0	1.4.0	1.3.1

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.219	0.219	0.215
Spark	2.4.3	2.4.2	2.4.2	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.12.0	1.12.0	1.12.0
Tez	0.9.2	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.13	3.4.13	3.4.13

Notas de la versión 5.25.0

Las siguientes notas de la versión incluyen información sobre la versión 5.25.0 de Amazon EMR. Los cambios son respecto a la versión 5.24.1.

Fecha de lanzamiento inicial: 17 de julio de 2019

Fecha de la última actualización: 30 de octubre de 2019

Amazon EMR 5.25.0

Actualizaciones

- AWS SDK para Java 1.11.566

- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- ZooKeeper 3.4.14

Nuevas características

- (30 de octubre de 2019) A partir de la versión 5.25.0 de Amazon EMR, podrá conectarse a la IU del servidor del historial de Spark desde la página Resumen del clúster o desde la pestaña Historial de aplicaciones de la consola. En lugar de configurar un proxy web a través de una conexión SSH, puede acceder rápidamente a la interfaz de usuario del servidor de historial de Spark para ver las métricas de las aplicaciones y acceder a los archivos de registro relevantes de los clústeres activos y terminados. Para más información, consulte Acceso fuera del clúster a las interfaces de usuario de aplicaciones persistentes en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha mejorado el rendimiento de algunas uniones mediante el uso de filtros de Bloom para hacer un filtrado previo de las entradas. La optimización está deshabilitada de forma predeterminada y se puede habilitar configurando el parámetro `spark.sql.bloomFilterJoin.enabled` de configuración de Spark en `true`.
  - Se ha mejorado el rendimiento de la agrupación de columnas por tipo de cadena.
  - Se mejoró la configuración predeterminada de memoria y núcleos del ejecutor Spark de los tipos de instancias R4 para clústeres sin instalación. HBase
  - Se ha resuelto un problema anterior relacionado con la característica de eliminación dinámica de particiones, por el que la tabla eliminada tenía que estar en el lado izquierdo de la unión.
  - Se ha mejorado la optimización de DISTINCT antes de INTERSECT para aplicarla a casos adicionales relacionados con alias.
  - Se ha mejorado la inferencia estadística del plan de SQL para JOIN seguida de consultas DISTINCT. Esta mejora se ha desactivado de forma predeterminada y se puede habilitar configurando el parámetro de configuración de Spark

`spark.sql.statsImprovements.enabled` como `true`. La característica Distinct antes de Intersect requiere esta optimización y se habilitará automáticamente cuando `spark.sql.optimizer.distinctBeforeIntersect.enabled` se establezca en `true`.

- Se ha optimizado el orden de unión con base en el tamaño de la tabla y de los filtros. La optimización se ha desactivado de forma predeterminada y se puede habilitar configurando el parámetro de configuración de Spark `spark.sql.optimizer.sizeBasedJoinReorder.enabled` en `true`.

Para más información, consulte Optimización del rendimiento de Spark.

- EMRFS
  - La configuración de EMRFS, `fs.s3.buckets.create.enabled`, ahora está deshabilitada de forma predeterminada. Tras realizar pruebas, se ha descubierto que al deshabilitar esta configuración se mejora el rendimiento y se evita la creación involuntaria de buckets de S3. Si su aplicación se basa en esta funcionalidad, puede habilitarla estableciendo la propiedad `fs.s3.buckets.create.enabled` en `true` en la clasificación de configuración `emrfs-site`. Para obtener información, consulte Suministro de una configuración durante la creación de un clúster.
- Mejoras en el cifrado de disco local y en el cifrado de S3 en las configuraciones de seguridad (5 de agosto de 2019)
  - Se ha separado la configuración del cifrado de Amazon S3 de la configuración del cifrado de disco local en la configuración de seguridad.
  - Se ha agregado una opción para habilitar el cifrado de EBS en la versión 5.24.0 y posteriores. Al seleccionar esta opción, se cifra el volumen del dispositivo raíz además de los volúmenes de almacenamiento. Las versiones anteriores requerían el uso de una AMI personalizada para cifrar el volumen del dispositivo raíz.
  - Para más información, consulte Opciones de cifrado en la Guía de administración de Amazon EMR.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del

periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.25.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres

veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.4	SDK de Amazon SageMaker Spark
emr-ddb	4.10.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.9.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR
emrfs	2.34.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.8.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-4	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.8.5-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.9	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.9	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.9	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.9	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.9	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.5-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.5-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.5-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.5-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.5-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.5-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.5-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.6	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.4.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	4.14.1- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.1- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.220	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.220	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.3	Clientes de línea de comando de Spark.
spark-history-server	2.4.3	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.3	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.

Componente	Versión	Descripción
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.13.1	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.2	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.14	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.14	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.25.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.25.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase

Clasificaciones	Descripción
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.

Clasificaciones	Descripción
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.

Clasificaciones	Descripción
recordserver-env	Cambie los valores en el entorno EMR. RecordServer
recordserver-conf	Cambie los valores en el archivo erver.properties RecordServer de EMR.
recordserver-log4j	Cambie los valores en el archivo log4j.properties RecordServer de EMR.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de Sqoop OraOop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.

Clasificaciones	Descripción
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.24.1

Versiones de las aplicaciones de la versión 5.24.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
AWS SDK for Java	1.11.546	1.11,546	1.11,519	1.11,519
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.4	0.9.4
Livy	0.6.0	0.6.0	0.5.0	0.5.0
MXNet	1.4.0	1.4.0	1.3.1	1.3.1

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.219	0.215	0.215
Spark	2.4.2	2.4.2	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.24.1

Las siguientes notas de la versión incluyen información sobre la versión 5.24.1 de Amazon EMR. Los cambios son respecto a la versión 5.24.0.

Fecha de lanzamiento inicial: 26 de junio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.24.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que

pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.9.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.8.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR
emrfs	2.33.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.8.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.

Componente	Versión	Descripción
hadoop-hdfs-journalnode	2.8.5-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.9	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.9	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.9	HBase cliente de línea de comandos.
hbase-rest-server	1.4.9	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.9	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.4-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-2	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.4-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.6	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.4.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.1- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.1- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.219	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.219	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.2	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.4.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.24.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.24.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.

Clasificaciones	Descripción
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site

Clasificaciones	Descripción
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.

Clasificaciones	Descripción
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.

Clasificaciones	Descripción
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.pr operties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.

Clasificaciones	Descripción
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.24.0

Versiones de las aplicaciones de la versión 5.24.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
AWS SDK for Java	1.11.546	1.11,519	1.11,519	1.11.510
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.7.1	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.4	0.9.4	0.9.4
Livy	0.6.0	0.5.0	0.5.0	0.5.0
MXNet	1.4.0	1.3.1	1.3.1	1.3.1

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.215	0.215	0.215
Spark	2.4.2	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.24.0

Las siguientes notas de la versión incluyen información sobre la versión 5.24.0 de Amazon EMR. Los cambios son respecto a la versión 5.23.0.

Fecha de lanzamiento inicial: 11 de junio de 2019

Fecha de la última actualización: 5 de agosto de 2019

Actualizaciones

- Flink 1.8.0
- Hue
- JupyterHub 0.9.6

- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK para Java 1.11,546
- Conectores y controladores:
  - Conector de DynamoDB 4.9.0
  - MariaDB Connector 2.4.1
  - Controlador JDBC de Amazon Redshift, versión 1.2.27.1051

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha agregado una optimización para eliminar las particiones de forma dinámica. La optimización se deshabilita de forma predeterminada. Para habilitarla, defina el parámetro de configuración de Spark `spark.sql.dynamicPartitionPruning.enabled` en `true`.
  - Se ha mejorado el rendimiento de las consultas `INTERSECT`. Esta optimización se deshabilita de forma predeterminada. Para habilitarla, defina el parámetro de configuración de Spark `spark.sql.optimizer.distinctBeforeIntersect.enabled` en `true`.
  - Se ha agregado una optimización para aplanar las subconsultas escalares con agregados que utilicen la misma relación. La optimización se deshabilita de forma predeterminada. Para habilitarla, defina el parámetro de configuración de Spark `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` en `true`.
  - Se ha mejorado la generación de código en todas las etapas.

  Para más información, consulte Optimización del rendimiento de Spark.
- Mejoras en el cifrado de disco local y en el cifrado de S3 en las configuraciones de seguridad (5 de agosto de 2019)
  - Se ha separado la configuración del cifrado de Amazon S3 de la configuración del cifrado de disco local en la configuración de seguridad.
  - Se ha agregado una opción para habilitar el cifrado de EBS. Al seleccionar esta opción, se cifra el volumen del dispositivo raíz además de los volúmenes de almacenamiento. Las versiones anteriores requerían el uso de una AMI personalizada para cifrar el volumen del dispositivo raíz.

- Para más información, consulte Opciones de cifrado en la Guía de administración de Amazon EMR.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

    > **Note**
    >
    > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.24.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.9.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.8.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.3.0	Conector S3Select de EMR

Componente	Versión	Descripción
emrfs	2.33.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.8.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-4	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.5-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-4	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.4.9	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.9	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.9	HBase cliente de línea de comandos.
hbase-rest-server	1.4.9	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.9	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.4-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-2	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.4-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.4.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.6	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.6.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.4.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.1- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.1- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.219	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.219	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico

Componente	Versión	Descripción
spark-client	2.4.2	Clientes de línea de comando de Spark.
spark-history-server	2.4.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.24.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.24.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.

Clasificaciones	Descripción
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.

Clasificaciones	Descripción
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.

Clasificaciones	Descripción
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.

Clasificaciones	Descripción
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.

Clasificaciones	Descripción
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.23.1

Versiones de las aplicaciones de la versión 5.23.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
AWS SDK for Java	1.11.519	1.11,519	1.11.510	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.1	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.23.1

Esta es una versión de revisión. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

Estado de soporte de Instance Metadata Service (IMDS) V2: los componentes de Amazon EMR 5.23.1, 5.27.1 y 5.32 o versiones posteriores se utilizan para todas las llamadas de IMDS. IMDSv2 Para las llamadas de IMDS en el código de su aplicación, puede usar ambos o configurar el IMDS

para que se use solo con el IMDSv1 fin de IMDSv2 aumentar la seguridad. IMDSv2 En otras versiones de EMR 5.x, la IMDSv1 desactivación provoca un error al iniciar el clúster.

Versiones de los componentes de la versión 5.23.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.8.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.7.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.

Componente	Versión	Descripción
emr-s3-select	1.2.0	Conector S3Select de EMR
emrfs	2.32.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.7.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.8.5-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-3	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.

Componente	Versión	Descripción
hadoop-yarn-timeline-server	2.8.5-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.9	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.9	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.9	HBase cliente de línea de comandos.
hbase-rest-server	1.4.9	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.9	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.

Componente	Versión	Descripción
hive-client	2.3.4-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.

Componente	Versión	Descripción
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.1- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.1- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.215	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.215	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.0	Clientes de línea de comando de Spark.
spark-history-server	2.4.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.23.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.23.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.

Clasificaciones	Descripción
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.

Clasificaciones	Descripción
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.

Clasificaciones	Descripción
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.

Clasificaciones	Descripción
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.

Clasificaciones	Descripción
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.23.0

Versiones de las aplicaciones de la versión 5.23.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
AWS SDK for Java	1.11.519	1.11.510	1.11.479	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.23.0

Las siguientes notas de la versión incluyen información sobre la versión 5.23.0 de Amazon EMR. Los cambios son respecto a la versión 5.22.0.

Fecha de lanzamiento inicial: 1 de abril de 2019

Fecha de la última actualización: 30 de abril de 2019

Actualizaciones

- AWS SDK para Java 1.11,519

Nuevas características

- (30 de abril de 2019) Con Amazon EMR 5.23.0 y versiones posteriores, puede lanzar un clúster con tres nodos principales para admitir la alta disponibilidad de aplicaciones como YARN Resource Manager, HDFS, Spark NameNode, Hive y Ganglia. El nodo principal ya no es un posible punto único de error con esta característica. Si se produce un error en uno de los nodos principales, se produce un error automáticamente en Amazon EMR a través de un nodo principal en espera y se sustituye el nodo principal con error por uno nuevo con las mismas acciones de arranque y configuración. Para más información, consulte Planificación y configuración de nodos principales.

Problemas conocidos

- Interfaz de usuario de Tez (corregida en la versión 5.26.0 de Amazon EMR)

  La IU de Tez no funciona en un clúster de EMR con varios nodos principales.
- Hue (corregido en la versión 5.24.0 de Amazon EMR)
  - Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Para evitar que aparezca el mensaje de error de Solr:

    1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
    2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

```
appblacklist = search
```

4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

```
sudo stop hue; sudo start hue
```

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.23.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.8.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.7.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.2.0	Conector S3Select de EMR

Componente	Versión	Descripción
emrfs	2.32.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.7.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-3	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.5-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-hdfs-journalnode	2.8.5-amzn-3	Servicio de HDFS para administrar los archivos de Hadoop periódico en clústeres de alta disponibilidad.
hadoop-httpfs-server	2.8.5-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.4.9	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.9	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.9	HBase cliente de línea de comandos.
hbase-rest-server	1.4.9	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.9	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.4-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-1	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.4-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.1- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.1- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.215	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.215	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico

Componente	Versión	Descripción
spark-client	2.4.0	Clientes de línea de comando de Spark.
spark-history-server	2.4.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.23.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.23.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.

Clasificaciones	Descripción
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.

Clasificaciones	Descripción
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.

Clasificaciones	Descripción
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.

Clasificaciones	Descripción
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.

Clasificaciones	Descripción
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.22.0

Versiones de las aplicaciones de la versión 5.22.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
AWS SDK for Java	1.11.510	1.11.479	1.11.479	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.0	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.22.0

Las siguientes notas de la versión incluyen información sobre la versión 5.22.0 de Amazon EMR. Los cambios son respecto a la versión 5.21.0.

> **Important**
>
> A partir de la versión 5.22.0 de Amazon EMR, Amazon EMR AWS utiliza Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Las versiones anteriores de Amazon EMR utilizan AWS Signature Version 2 en algunos casos, a menos que las notas de la versión indiquen que se utiliza exclusivamente Signature Version 4. Para obtener más información, consulte Autenticación de solicitudes (versión de AWS firma 4) y Autenticación de solicitudes (versión de AWS firma 2) en la Guía para desarrolladores de Amazon Simple Storage Service.

Fecha de lanzamiento inicial: 20 de marzo de 2019

Actualizaciones

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- Conectores y controladores:
  - Conector de DynamoDB 4.8.0
  - MariaDB Connector 2.2.6
  - Controlador JDBC de Amazon Redshift, versión 1.2.20.1043

Nuevas características

- Se modificó la configuración de EBS predeterminada para los tipos de EC2 instancias con almacenamiento exclusivo para EBS. Al crear un clúster utilizando la versión 5.22.0 y posteriores de Amazon EMR, la cantidad predeterminada de almacenamiento de EBS aumenta en función del tamaño de la instancia. Además, podemos dividir un mayor almacenamiento en varios volúmenes, lo que ofrece un mayor desempeño de IOPS. Si desea utilizar una configuración de almacenamiento para instancias de EBS diferente, puede especificarlo al crear un clúster de EMR o añadir nodos a un clúster existente. Para más información sobre la cantidad de almacenamiento y la cantidad de volúmenes asignados de forma predeterminada a cada tipo de instancia, consulte

Almacenamiento predeterminado de EBS para instancias en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha introducido una nueva propiedad de configuración para Spark en YARN, `spark.yarn.executor.memoryOverheadFactor`. El valor de esta propiedad es un factor de escala que establece el valor de la sobrecarga de memoria en un porcentaje de la memoria del ejecutor, con un mínimo de 384 MB. Si la sobrecarga de memoria se establece de forma explícita mediante `spark.yarn.executor.memoryOverhead`, esta propiedad no tendrá ningún efecto. El valor predeterminado es `0.1875`, que representa el 18,75 %. Este valor predeterminado para Amazon EMR deja más espacio en los contenedores de YARN para la sobrecarga de memoria del ejecutor que el 10 % predeterminado que ha establecido Spark internamente. El valor predeterminado de Amazon EMR del 18,75 % mostró empíricamente menos errores relacionados con la memoria en los puntos de referencia de TPC-DS.
  - Se ha agregado portabilidad con versiones anteriores para SPARK-26316 a fin de mejorar el rendimiento.
- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.

Problemas conocidos

- Hue (corregido en la versión 5.24.0 de Amazon EMR)
  - Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    ```

```
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Para evitar que aparezca el mensaje de error de Solr:

1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

   ```
   sudo vim /etc/hue/conf/hue.ini
   ```

3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

   ```
   appblacklist = search
   ```

4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

   ```
   sudo stop hue; sudo start hue
   ```

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.22.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.8.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.6.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.2.0	Conector S3Select de EMR
emrfs	2.31.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.7.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.8.5-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.9	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.9	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.9	HBase cliente de línea de comandos.
hbase-rest-server	1.4.9	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.9	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.4-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.1.0	Cliente de línea de comando de Oozie.
oozie-server	5.1.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.1- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.1- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.215	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.215	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.0	Clientes de línea de comando de Spark.
spark-history-server	2.4.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.22.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.22.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.

Clasificaciones	Descripción
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log

Clasificaciones	Descripción
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.

Clasificaciones	Descripción
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.

Clasificaciones	Descripción
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.21.2

Versiones de las aplicaciones de la versión 5.21.2

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
AWS SDK for Java	1.11.479	1.11.479	1.11.479	1.11.461
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.7.0	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.21.2

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.21.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.5.1	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.2.0	Conector S3Select de EMR
emrfs	2.30.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.7.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.8.5-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.8	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.8	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.8	HBase cliente de línea de comandos.
hbase-rest-server	1.4.8	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.8	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.4-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4,14,0- -1,4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.215	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.215	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.0	Clientes de línea de comando de Spark.
spark-history-server	2.4.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.21.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.21.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.

Clasificaciones	Descripción
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log

Clasificaciones	Descripción
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.

Clasificaciones	Descripción
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.

Clasificaciones	Descripción
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.21.1

Versiones de las aplicaciones de la versión 5.21.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
AWS SDK for Java	1.11.479	1.11.479	1.11.461	1.11.461
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.6.2	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.214	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.21.1

Las siguientes notas de la versión incluyen información sobre la versión 5.21.1 de Amazon EMR. Los cambios son respecto a la versión 5.21.0.

Fecha de lanzamiento inicial: 18 de julio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.21.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.1	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.2.0	Conector S3Select de EMR
emrfs	2.30.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.7.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.5-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.8.5-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.8	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.8	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.8	HBase cliente de línea de comandos.
hbase-rest-server	1.4.8	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.8	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.4-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.215	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.215	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.0	Clientes de línea de comando de Spark.
spark-history-server	2.4.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.21.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.21.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.

Clasificaciones	Descripción
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log

Clasificaciones	Descripción
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.

Clasificaciones	Descripción
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.

Clasificaciones	Descripción
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.21.0

Versiones de las aplicaciones de la versión 5.21.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
AWS SDK for Java	1.11.479	1.11.461	1.11.461	1.11.433
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.4
Scala	2.11.12	2.11.12	2.11.12	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.6.2	1.6.2	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.7
HCatalog	2.3.4	2.3.4	2.3.4	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.3

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.214	0.214	0.212
Spark	2.4.0	2.4.0	2.4.0	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.11.0
Tez	0.9.1	0.9.1	0.9.1	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.21.0

Las siguientes notas de la versión incluyen información sobre la versión 5.21.0 de Amazon EMR. Los cambios son respecto a la versión 5.20.0.

Fecha de lanzamiento inicial: 18 de febrero de 2019

Fecha de la última actualización: 3 de abril de 2019

Actualizaciones

- Flink 1.7.0
- Presto 0.215
- AWS SDK para Java 1.11.479

Nuevas características

- (3 de abril de 2019) Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Cambios, mejoras y problemas resueltos

- Zeppelin
  - Se ha adaptado ZEPPELIN-3878.

Problemas conocidos

- Hue (corregido en la versión 5.24.0 de Amazon EMR)
  - Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    ```

```
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Para evitar que aparezca el mensaje de error de Solr:

1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

   ```
   sudo vim /etc/hue/conf/hue.ini
   ```
3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

   ```
   appblacklist = search
   ```
4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

   ```
   sudo stop hue; sudo start hue
   ```

- Tez
  - Este problema se ha solucionado en la versión 5.22.0 de Amazon EMR.

    Cuando se conecta a la interfaz de usuario de Tez en http: //:8080/tez-ui *`MasterDNS`* a través de una conexión SSH con el nodo principal del clúster, aparece el error «Falló la operación del adaptador; el servidor Timeline (ATS) está fuera de su alcance». Either it is down, or CORS is not enabled" o bien las tareas muestran inesperadamente N/A.

    Esto se debe a que la interfaz de usuario de Tez realiza solicitudes al servidor YARN Timeline con `localhost` en vez de con el nombre de host del nodo principal. Como solución alternativa, hay un script disponible para ejecutar como acción o paso de arranque. El script actualiza el nombre del host en el archivo `configs.env` de Tez. Para más información y obtener la ubicación del script, consulte Bootstrap Instructions.
- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.21.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.1	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.11.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.2.0	Conector S3Select de EMR
emrfs	2.30.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.7.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del

Componente	Versión	Descripción
		ecosistema de Hadoop junto con el agente de monitoriz ación de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.8.5-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.8	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.8	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.8	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.8	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.8	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.215	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.215	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.0	Clientes de línea de comando de Spark.
spark-history-server	2.4.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.

Componente	Versión	Descripción
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.21.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.21.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase

Clasificaciones	Descripción
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.

Clasificaciones	Descripción
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.20.1

Versiones de las aplicaciones de la versión 5.20.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
AWS SDK for Java	1.11.461	1.11.461	1.11.433	1.11.433
Python	2.7, 3.6	2.7, 3.6	2.7, 3.4	2.7, 3.4
Scala	2.11.12	2.11.12	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
Delta	-	-	-	-
Flink	1.6.2	1.6.2	1.6.1	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.7	1.4.7
HCatalog	2.3.4	2.3.4	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.0	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.214	0.212	0.212
Spark	2.4.0	2.4.0	2.3.2	2.3.2

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.11.0	1.11.0
Tez	0.9.1	0.9.1	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

Notas de la versión 5.20.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.20.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.1	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.2.0	Conector S3Select de EMR
emrfs	2.29.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.6.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del

Componente	Versión	Descripción
		ecosistema de Hadoop junto con el agente de monitoriz ación de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.8.5-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.8	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.8	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.8	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.8	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.8	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.4-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.

Componente	Versión	Descripción
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.214	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.214	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.0	Clientes de línea de comando de Spark.
spark-history-server	2.4.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.

Componente	Versión	Descripción
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.20.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.20.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase

Clasificaciones	Descripción
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.

Clasificaciones	Descripción
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.20.0

Versiones de las aplicaciones de la versión 5.20.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
AWS SDK for Java	1.11.461	1.11.433	1.11.433	1.11.393
Python	2.7, 3.6	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.12	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
Delta	-	-	-	-
Flink	1.6.2	1.6.1	1.6.1	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.7	1.4.7	1.4.7
HCatalog	2.3.4	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.4
Hive	2.3.4	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.0	1.3.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.212	0.212	0.210
Spark	2.4.0	2.3.2	2.3.2	2.3.2

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.11.0	1.11.0	1.9.0
Tez	0.9.1	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.12

Notas de la versión 5.20.0

Las siguientes notas de la versión incluyen información sobre la versión 5.20.0 de Amazon EMR. Los cambios son respecto a la versión 5.19.0.

Fecha de lanzamiento inicial: 18 de diciembre de 2018

Fecha de la última actualización: 22 de enero de 2019

Actualizaciones

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK para Java 1.11.461

Nuevas características

- (22 de enero de 2019) Se ha mejorado Kerberos en Amazon EMR para admitir la autenticación de las entidades principales desde un KDC externo. Esto centraliza la administración de entidades principales porque varios clústeres pueden compartir un único KDC externo. Además, el KDC externo puede tener una relación de confianza entre ámbitos con un dominio de Active Directory. Esto permite que todos los clústeres autentiquen entidades principales desde Active Directory. Para más información, consulte Uso de la autenticación de Kerberos en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- AMI de Amazon Linux predeterminada para Amazon EMR
  - Se ha actualizado el paquete de Python3 de Python 3.4 a 3.6.
- El confirmador optimizado para S3 de EMRFS
  - El confirmador optimizado para S3 de EMRFS ahora está habilitado de forma predeterminada, lo que mejora el rendimiento de escritura. Para obtener más información, consulte Uso del confirmador optimizado para S3 de EMRFS.
- Hive
  - Se ha adaptado HIVE-16686.
- Glue con Spark y Hive
  - En EMR 5.20.0 o versiones posteriores, la reducción de particiones en paralelo se habilita automáticamente para Spark y Hive cuando se utiliza AWS Glue Data Catalog como metaalmacén. Este cambio reduce significativamente el tiempo de planificación de consultas al ejecutar varias solicitudes en paralelo para recuperar particiones. El número total de segmentos que se pueden ejecutar simultáneamente oscila entre 1 y 10. El valor predeterminado es 5, que es la configuración recomendada. Para cambiarlo, especifique la propiedad `aws.glue.partition.num.segments` en la clasificación de la configuración `hive-site`. Si se produce una limitación, puede desactivar la característica cambiando el valor a 1. Para más información, consulte Estructura Segment de AWS Glue.

Problemas conocidos

- Hue (corregido en la versión 5.24.0 de Amazon EMR)

- Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

  ```
  Solr server could not be contacted properly:
  HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
  port=1978): Max retries exceeded with url: /solr/admin/info/
  system?user.name=hue&doAs=administrator&wt=json (Caused by
  NewConnectionError(': Failed to establish a new connection: [Errno 111]
  Connection refused',))
  ```

  Para evitar que aparezca el mensaje de error de Solr:

  1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
  2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

     ```
     sudo vim /etc/hue/conf/hue.ini
     ```
  3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

     ```
     appblacklist = search
     ```
  4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

     ```
     sudo stop hue; sudo start hue
     ```

- Tez
  - Este problema se ha solucionado en la versión 5.22.0 de Amazon EMR.

    Cuando te conectas a la interfaz de usuario de Tez en http: //:8080/tez-ui *MasterDNS* mediante una conexión SSH con el nodo principal del clúster, aparece el error «Falló la operación del adaptador; el servidor Timeline (ATS) está fuera de alcance». Either it is down, or CORS is not enabled" o bien las tareas muestran inesperadamente N/A.

    Esto se debe a que la interfaz de usuario de Tez realiza solicitudes al servidor YARN Timeline con `localhost` en vez de con el nombre de host del nodo principal. Como solución alternativa, hay un script disponible para ejecutar como acción o paso de arranque. El script actualiza el nombre del host en el archivo `configs.env` de Tez. Para más información y obtener la ubicación del script, consulte Bootstrap Instructions.

- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versiones de los componentes de la versión 5.20.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.1	SDK de Amazon SageMaker Spark
emr-ddb	4.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.1	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.2.0	Conector S3Select de EMR

Componente	Versión	Descripción
emrfs	2.29.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.6.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.5-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.8	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.8	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.8	HBase cliente de línea de comandos.
hbase-rest-server	1.4.8	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.8	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.4-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.4-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.4-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.4-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.4-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.4-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.4-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.3.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.1	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.214	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.214	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.4.0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.4.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.4.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.4.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.12.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.9.1	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.20.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.20.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.

Clasificaciones	Descripción
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site

Clasificaciones	Descripción
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.

Clasificaciones	Descripción
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.

Clasificaciones	Descripción
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.

Clasificaciones	Descripción
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.19.1

Versiones de las aplicaciones de la versión 5.19.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
AWS SDK for Java	1.11.433	1.11.433	1.11.393	1.11.393
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.1	1.6.0	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.7
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.3.0	1.2.0	1.2.0

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.212	0.210	0.210
Spark	2.3.2	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.11.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.12	3.4.12

Notas de la versión 5.19.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.19.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.0	SDK de Amazon SageMaker Spark
emr-ddb	4.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.1	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.1.0	Conector S3Select de EMR

Componente	Versión	Descripción
emrfs	2.28.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.6.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-0	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.5-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.5-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.7	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.7	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.7	HBase cliente de línea de comandos.
hbase-rest-server	1.4.7	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.7	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-2	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.3-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.212	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.212	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.2	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.3.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.19.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.19.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.

Clasificaciones	Descripción
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site

Clasificaciones	Descripción
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.

Clasificaciones	Descripción
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.

Clasificaciones	Descripción
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.

Clasificaciones	Descripción
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.19.0

Versiones de las aplicaciones de la versión 5.19.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
AWS SDK for Java	1.11.433	1.11.393	1.11.393	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.0	1.6.0	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.2.0	1.2.0	1.2.0

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.210	0.210	0.206
Spark	2.3.2	2.3.2	2.3.2	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.7.3
ZooKeeper	3.4.13	3.4.12	3.4.12	3.4.12

Notas de la versión 5.19.0

Las siguientes notas de la versión incluyen información sobre la versión 5.19.0 de Amazon EMR. Los cambios son respecto a la versión 5.18.0.

Fecha de lanzamiento inicial: 7 de noviembre de 2018

Fecha de la última actualización: 19 de noviembre de 2018

Actualizaciones

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4

- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- ZooKeeper 3.4.13
- AWS SDK para Java 1.11.433

Nuevas características

- (19 de noviembre de 2018) Cuadernos de EMR es un entorno administrado basado en cuadernos de Jupyter. Es compatible con los núcleos mágicos de Spark para PySpark Spark SQL, Spark R y Scala. Cuadernos de EMR se puede utilizar con clústeres creados con la versión 5.18.0 y posteriores de Amazon EMR. Para más información, consulte Uso de Cuadernos de EMR en la Guía de administración de Amazon EMR.
- El confirmador optimizado para S3 de EMRFS está disponible al escribir archivos Parquet con Spark y EMRFS. Este confirmador mejora el rendimiento de escritura. Para obtener más información, consulte Uso del confirmador optimizado para S3 de EMRFS.

Cambios, mejoras y problemas resueltos

- YARN
  - Se ha modificado la lógica que limita el proceso maestro de la aplicación a ejecutarse en los nodos principales. Esta funcionalidad ahora utiliza la característica y las propiedades de las etiquetas de nodos de YARN en las clasificaciones de configuración `yarn-site` y `capacity-scheduler`. Para obtener información, consulte https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.
- AMI de Amazon Linux predeterminada para Amazon EMR
  - `ruby18`, `php56` y `gcc48` ya no se instalan de forma predeterminada. Se pueden instalar si se desea utilizando `yum`.
  - La gema aws-sdk ruby ya no se instala de forma predeterminada. Si lo desea, se puede instalar utilizando `gem install aws-sdk`. También se pueden instalar componentes específicos. Por ejemplo, `gem install aws-sdk-s3`.

Problemas conocidos

- Cuadernos de EMR: en algunas circunstancias, con varios editores de cuadernos abiertos, puede parecer que el editor de cuadernos no puede conectarse al clúster. Si esto ocurre, borre las cookies del navegador y, a continuación, vuelva a abrir los editores de cuadernos.
- CloudWatch ContainerPending Escalado métrico y automático: (corregido en la versión 5.20.0) Amazon EMR puede emitir un valor negativo para. `ContainerPending` Si `ContainerPending` se utiliza en una regla de escalado automático, el escalado automático no se comportará como se esperaba. Evite usar `ContainerPending` con el escalado automático.
- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.

Versiones de los componentes de la versión 5.19.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.2.0	SDK de Amazon SageMaker Spark

Componente	Versión	Descripción
emr-ddb	4.7.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.1	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.1.0	Conector S3Select de EMR
emrfs	2.28.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.6.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.5-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.5-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.5-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.5-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.5-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.5-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.5-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.5-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.5-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.5-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.7	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.7	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.7	HBase cliente de línea de comandos.
hbase-rest-server	1.4.7	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.7	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.3-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.3-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.9.4	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.3.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.212	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.212	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.2	Clientes de línea de comando de Spark.
spark-history-server	2.3.2	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.11.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.13	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.13	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.19.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.19.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.

Clasificaciones	Descripción
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log

Clasificaciones	Descripción
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.

Clasificaciones	Descripción
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.

Clasificaciones	Descripción
presto-connector-memory	Cambiar los valores del archivo memory.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
presto-connector-tpcds	Cambiar los valores del archivo tpcds.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.18.1

Versiones de las aplicaciones de la versión 5.18.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
AWS SDK for Java	1.11.393	1.11.393	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.6.0	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.210	0.206	0.206
Spark	2.3.2	2.3.2	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas de la versión 5.18.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.18.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.1.3	SDK de Amazon SageMaker Spark
emr-ddb	4.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.5.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.1.0	Conector S3Select de EMR
emrfs	2.27.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.6.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.8.4-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.4-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.4-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.4-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.4-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.4-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.7	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.7	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.7	HBase cliente de línea de comandos.
hbase-rest-server	1.4.7	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.7	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.3-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.3-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.2.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.210	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.210	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.2	Clientes de línea de comando de Spark.
spark-history-server	2.3.2	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.

Componente	Versión	Descripción
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.18.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.18.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.

Clasificaciones	Descripción
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log

Clasificaciones	Descripción
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.

Clasificaciones	Descripción
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.

Clasificaciones	Descripción
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.18.0

Versiones de las aplicaciones de la versión 5.18.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
AWS SDK for Java	1.11.393	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.5.2	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.6	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.206	0.206	0.206
Spark	2.3.2	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas de la versión 5.18.0

Las siguientes notas de la versión incluyen información sobre la versión 5.18.0 de Amazon EMR. Los cambios son respecto a la versión 5.17.0.

Fecha de lanzamiento inicial: 24 de octubre de 2018

Actualizaciones

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

Nuevas características

- A partir de Amazon EMR 5.18.0, puede utilizar el repositorio de artefactos de Amazon EMR para compilar el código de trabajos con las versiones exactas de las bibliotecas y dependencias que están disponibles con determinadas versiones de Amazon EMR. Para obtener más información, consulte Comprobación de dependencias mediante el repositorio de artefactos de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Hive
  - Se ha agregado soporte para S3 Select. Para obtener más información, consulte Uso de S3 Select con Hive para mejorar el rendimiento.
- Presto
  - Se ha agregado soporte para S3 Select Pushdown. Para obtener más información, consulte Uso de S3 Select Pushdown con Presto para mejorar el rendimiento.
- Spark
  - Se ha cambiado la configuración predeterminada de log4j para Spark y ahora permite distribuir los registros de contenedores cada hora para los trabajos de streaming de Spark. Esto ayuda a evitar que se eliminen los registros de los trabajos de streaming de Spark de ejecución prolongada.

Versiones de los componentes de la versión 5.18.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de

Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.1.3	SDK de Amazon SageMaker Spark
emr-ddb	4.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.1.0	Conector S3Select de EMR
emrfs	2.27.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
flink-client	1.6.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.4-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.4-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.4-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.4-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.4-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.4-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.4-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.7	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.7	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.7	HBase cliente de línea de comandos.
hbase-rest-server	1.4.7	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.7	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.3-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-2	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.3-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
nginx	1.12.1	nginx [engine x] es un servidor HTTP y proxy inverso
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.2.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.210	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.210	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.2	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.3.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.8.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.18.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.18.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.

Clasificaciones	Descripción
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site

Clasificaciones	Descripción
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.

Clasificaciones	Descripción
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.

Clasificaciones	Descripción
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.

Clasificaciones	Descripción
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.17.2

Versiones de las aplicaciones de la versión 5.17.2

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.2	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.6	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.206	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas de la versión 5.17.2

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.17.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.1.3	SDK de Amazon SageMaker Spark
emr-ddb	4.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.0.0	Conector S3Select de EMR
emrfs	2.26.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
flink-client	1.5.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.4-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.4-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.4-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.4-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.4-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.4-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.4-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.6	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.6	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.6	HBase cliente de línea de comandos.
hbase-rest-server	1.4.6	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.6	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-1	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.3-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.2.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.0.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.206	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.206	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.1	Clientes de línea de comando de Spark.
spark-history-server	2.3.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.

Componente	Versión	Descripción
spark-on-yarn	2.3.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.17.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.17.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.

Clasificaciones	Descripción
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.

Clasificaciones	Descripción
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.17.1

Versiones de las aplicaciones de la versión 5.17.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.0	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.203	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas de la versión 5.17.1

Las siguientes notas de la versión incluyen información sobre la versión 5.17.1 de Amazon EMR. Los cambios son respecto a la versión 5.17.0.

Fecha de lanzamiento inicial: 18 de julio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Versiones de los componentes de la versión 5.17.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.1.3	SDK de Amazon SageMaker Spark
emr-ddb	4.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.0.0	Conector S3Select de EMR
emrfs	2.26.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.5.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del

Componente	Versión	Descripción
		ecosistema de Hadoop junto con el agente de monitoriz ación de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.4-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.4-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.4-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.4-amzn-1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.8.4-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.4-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.6	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.6	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.6	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.6	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.6	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.3-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.2.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.206	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.206	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.1	Clientes de línea de comando de Spark.
spark-history-server	2.3.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tensorflow	1.9.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.17.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.17.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.

Clasificaciones	Descripción
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.

Clasificaciones	Descripción
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.

Clasificaciones	Descripción
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.17.0

Versiones de las aplicaciones de la versión 5.17.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTensorFlow Tez, Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
AWS SDK for Java	1.11.336	1.11.336	1.11.336	1.11.333
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.0	1.5.0	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
HBase	1.4.6	1.4.4	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.4.0
MXNet	1.2.0	1.2.0	1.2.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.203	0.203	0.194
Spark	2.3.1	2.3.1	2.3.1	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas de la versión 5.17.0

Las siguientes notas de la versión incluyen información sobre la versión 5.17.0 de Amazon EMR. Los cambios son respecto a la versión 5.16.0.

Fecha de lanzamiento inicial: 30 de agosto de 2018

Actualizaciones

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

Nuevas características

- Se ha agregado compatibilidad con TensorFlow. Para obtener más información, consulte TensorFlow.

Cambios, mejoras y problemas resueltos

- JupyterHub
  - Se ha agregado compatibilidad con la persistencia de cuadernos en Amazon S3. Para obtener más información, consulte Configuración de la persistencia de los cuadernos en Amazon S3.
- Spark
  - Se ha agregado soporte para S3 Select. Para obtener más información, consulte Uso de S3 Select con Spark para mejorar el rendimiento de las consultas.
- Se han resuelto los problemas con las métricas de CloudWatch y la característica de escalado automático en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR.

Problemas conocidos

- Al crear un clúster que utiliza Kerberos con Livy instalado, Livy produce un error que indica que la autenticación simple no está habilitada. El problema se resuelve al reiniciar el servidor de Livy. Como solución alternativa, agregue un paso durante la creación del clúster que ejecuta `sudo restart livy-server` en el nodo principal.
- Si utiliza una AMI personalizada de Amazon Linux basada en una AMI de Amazon Linux con una fecha de creación del 11 de agosto de 2018, el servidor de Oozie no podrá iniciarse. Si utiliza Oozie, cree una AMI personalizada basada en un ID de AMI de Amazon Linux con una fecha de creación diferente. Puede utilizar el siguiente AWS CLI comando para obtener una lista de imágenes IDs para todas las HVM de Amazon Linux AMIs con una versión 2018.03, junto con la fecha de lanzamiento, de forma que pueda elegir una AMI de Amazon Linux adecuada como base. MyRegion Sustitúyalo por tu identificador de región, como us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

Versiones de los componentes de la versión 5.17.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.1.3	SDK de Amazon SageMaker Spark
emr-ddb	4.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.5.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emr-s3-select	1.0.0	Conector S3Select de EMR
emrfs	2.26.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.5.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.4-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.4-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.4-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.4-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.4-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.4-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.4-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.4-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.4-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.6	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.6	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.6	HBase cliente de línea de comandos.
hbase-rest-server	1.4.6	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.6	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.3-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.3-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.2.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.206	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.206	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.1	Clientes de línea de comando de Spark.
spark-history-server	2.3.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tensorflow	1.9.0	TensorFlow biblioteca de software de código abierto para computación numérica de alto rendimiento.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.

Componente	Versión	Descripción
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.17.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.17.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-s3-conf	Configurar la persistencia en S3 del bloc de notas de Jupyter.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.

Clasificaciones	Descripción
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.

Clasificaciones	Descripción
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.16.1

Versiones de las aplicaciones de la versión 5.16.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTez Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
AWS SDK for Java	1.11.336	1.11.336	1.11.333	1.11.333
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.5.0	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.4.0	0.4.0
MXNet	1.2.0	1.2.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.203	0.194	0.194
Spark	2.3.1	2.3.1	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

Notas de la versión 5.16.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.16.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.1.0	SDK de Amazon SageMaker Spark
emr-ddb	4.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.25.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.5.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.4-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.8.4-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.4-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.4-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.4-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.4-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.4-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.4-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.4-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.4.4	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.4	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.4	HBase cliente de línea de comandos.
hbase-rest-server	1.4.4	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.4	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.3-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-1	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.3-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.3-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.2.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.203	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.203	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.1	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.3.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.16.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.16.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.

Clasificaciones	Descripción
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.16.0

Versiones de las aplicaciones de la versión 5.16.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTez Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
AWS SDK for Java	1.11.336	1.11.333	1.11.333	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
HBase	1.4.4	1.4.4	1.4.4	1.4.2
HCatalog	2.3.3	2.3.3	2.3.3	2.3.2
Hadoop	2.8.4	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.4.0	0.4.0	0.4.0
MXNet	1.2.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	4.3.0
Phoenix	4.14.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.194	0.194	0.194
Spark	2.3.1	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.10

Notas de la versión 5.16.0

Las siguientes notas de la versión incluyen información sobre la versión 5.16.0 de Amazon EMR. Los cambios son respecto a la versión 5.15.0.

Fecha de lanzamiento inicial: 19 de julio de 2018

Actualizaciones

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK para Java 1.11.336
- CUDA 9.2
- Controlador JDBC de Amazon Redshift, versión 1.2 15.1025

Cambios, mejoras y problemas resueltos

- HBase
  - Se ha adaptado HBASE-20723
- Presto
  - Cambios de configuración para admitir la autenticación LDAP. Para obtener más información, consulte Uso de la autenticación LDAP para Presto en Amazon EMR.

- Spark
  - La versión 2.3.1 de Apache Spark, disponible a partir de la versión 5.16.0 de Amazon EMR, aborda CVE-2018-8024 y CVE-2018-1334. Se recomienda migrar las versiones anteriores de Spark a la versión 2.3.1 o posteriores de Spark.

Problemas conocidos

- Esta versión de lanzamiento no es compatible con los tipos de instancias c1.medium ni m1.small. Los clústeres que utilizan cualquiera de estos tipos de instancias no se inician. Como solución alternativa, especifique un tipo de instancia diferente o utilice una versión de lanzamiento diferente.
- Al crear un clúster que utiliza Kerberos con Livy instalado, Livy produce un error que indica que la autenticación simple no está habilitada. El problema se resuelve al reiniciar el servidor de Livy. Como solución alternativa, agregue un paso durante la creación del clúster que ejecuta `sudo restart livy-server` en el nodo principal.
- Una vez que se reinicie el nodo principal o se reinicie el controlador de instancias, no se recopilarán las CloudWatch métricas y la función de escalado automático no estará disponible en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR. Este problema se ha corregido en la versión 5.17.0 de Amazon EMR.

Versiones de los componentes de la versión 5.16.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.1.0	SDK de Amazon SageMaker Spark
emr-ddb	4.6.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.25.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.5.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.4-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.4-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.4-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.4-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.4-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.4-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.4-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.4-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.4-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.4-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.4	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.4	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.4	HBase cliente de línea de comandos.
hbase-rest-server	1.4.4	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.4	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.3-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.3-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.5.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.2.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.2.88	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.14.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.14.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.203	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.203	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.1	Clientes de línea de comando de Spark.
spark-history-server	2.3.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.16.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.16.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.

Clasificaciones	Descripción
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.

Clasificaciones	Descripción
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.

Clasificaciones	Descripción
presto-password-authenticator	Cambiar los valores en el archivo password-authenticator.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.

Clasificaciones	Descripción
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.

Clasificaciones	Descripción
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.15.1

Versiones de las aplicaciones de la versión 5.15.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTez Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
AWS SDK for Java	1.11.333	1.11.333	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.2	1.4.2
HCatalog	2.3.3	2.3.3	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.10	3.4.10

Notas de la versión 5.15.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.15.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.

Componente	Versión	Descripción
emrfs	2.24.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.4	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.4	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.4	HBase cliente de línea de comandos.
hbase-rest-server	1.4.4	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.4	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.3-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.3-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.3-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.1.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	5.0.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	5.0.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.194	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.194	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.0	Clientes de línea de comando de Spark.
spark-history-server	2.3.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.

Componente	Versión	Descripción
spark-on-yarn	2.3.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.15.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.15.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase

Clasificaciones	Descripción
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.

Clasificaciones	Descripción
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.

Clasificaciones	Descripción
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.15.0

Versiones de las aplicaciones de la versión 5.15.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTez Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
AWS SDK for Java	1.11.333	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.2	1.4.2	1.4.2
HCatalog	2.3.3	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.2	2.3.2	2.3.2

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
Hudi	-	-	-	-
Hue	4.2.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.10	3.4.10	3.4.10

Notas de la versión 5.15.0

Las siguientes notas de la versión incluyen información sobre la versión 5.15.0 de Amazon EMR. Los cambios son respecto a la versión 5.14.0.

Fecha de lanzamiento inicial: 21 de junio de 2018

Actualizaciones

- Actualizado HBase a 1.4.4
- Se ha actualizado Hive a la versión 2.3.3
- Se ha actualizado Hue a la versión 4.2.0
- Se ha actualizado Oozie a la versión 5.0.0
- Se ha actualizado ZooKeeper a la versión 3.4.12
- Se actualizó el AWS SDK a la versión 1.11.333

Cambios, mejoras y problemas resueltos

- Hive
  - Se ha adaptado HIVE-18069
- Hue
  - Se ha actualizado Hue para autenticarse correctamente con Livy cuando Kerberos está habilitado. Ahora se admite Livy cuando se usa Kerberos con Amazon EMR.
- JupyterHub
  - Se actualizó JupyterHub para que Amazon EMR instale las bibliotecas cliente LDAP de forma predeterminada.
  - Se ha corregido un error en el script que genera los certificados autofirmados.

Problemas conocidos

- Esta versión de lanzamiento no es compatible con los tipos de instancias c1.medium ni m1.small. Los clústeres que utilizan cualquiera de estos tipos de instancias no se inician. Como solución alternativa, especifique un tipo de instancia diferente o utilice una versión de lanzamiento diferente.
- Una vez que se reinicie el nodo principal o se reinicie el controlador de instancias, no se recopilarán las CloudWatch métricas y la función de escalado automático no estará disponible

en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR. Este problema se ha corregido en la versión 5.17.0 de Amazon EMR.

Versiones de los componentes de la versión 5.15.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.24.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.8.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.4.4	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.4	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.4	HBase cliente de línea de comandos.
hbase-rest-server	1.4.4	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.4	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.3-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.3-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.3-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.3-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.3-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.3-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.3-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.2.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.1.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	5.0.0	Cliente de línea de comando de Oozie.
oozie-server	5.0.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.194	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.194	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.3.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.12	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.12	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.15.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.15.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.14.2

Versiones de las aplicaciones de la versión 5.14.2

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTez Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.14.2

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.14.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.23.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.2	HBase cliente de línea de comandos.
hbase-rest-server	1.4.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.2-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.1.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.194	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.

Componente	Versión	Descripción
presto-worker	0.194	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.0	Clientes de línea de comando de Spark.
spark-history-server	2.3.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.14.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.14.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.

Clasificaciones	Descripción
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.

Clasificaciones	Descripción
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.

Clasificaciones	Descripción
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.

Clasificaciones	Descripción
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.14.1

Versiones de las aplicaciones de la versión 5.14.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTez Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.14.1

Las siguientes notas de la versión incluyen información sobre la versión 5.14.1 de Amazon EMR. Los cambios son respecto a la versión 5.14.0.

Fecha de lanzamiento inicial: 17 de octubre de 2018

Se ha actualizado la AMI predeterminada de Amazon EMR para abordar posibles vulnerabilidades de seguridad.

Versiones de los componentes de la versión 5.14.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.23.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto

Componente	Versión	Descripción
		con el agente de monitoriz ación de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.8.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.2	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.4.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.1.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.194	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.194	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.0	Clientes de línea de comando de Spark.
spark-history-server	2.3.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.

Componente	Versión	Descripción
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.14.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.14.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.

Clasificaciones	Descripción
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log

Clasificaciones	Descripción
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_notebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.

Clasificaciones	Descripción
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.

Clasificaciones	Descripción
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.

Clasificaciones	Descripción
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.14.0

Versiones de las aplicaciones de la versión 5.14.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog,, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, SparkSqoopTez Zeppelin, y ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR

- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
AWS SDK for Java	1.11.297	1.11.297	1.11.297	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	-	-	-

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.188
Spark	2.3.0	2.3.0	2.3.0	2.2.1
Sqoop	1.4.7	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.14.0

Las siguientes notas de la versión incluyen información sobre la versión 5.14.0 de Amazon EMR. Los cambios son respecto a la versión 5.13.0.

Fecha de lanzamiento inicial: 4 de junio de 2018

Actualizaciones

- Se ha actualizado Flink a la versión 1.4.2

- Se actualizó Apache MXnet a la versión 1.1.0
- Se ha actualizado Apache Sqoop a la versión 1.4.7

Nuevas características

- Se agregó JupyterHub soporte. Para obtener más información, consulte JupyterHub.

Cambios, mejoras y problemas resueltos

- EMRFS
  - La cadena userAgent de las solicitudes para Amazon S3 se ha actualizado para incluir la información de usuario y de grupo de la entidad principal que la invoca. Esto se puede usar con AWS CloudTrail los registros para un seguimiento más completo de las solicitudes.
- HBase
  - Se ha incluido HBASE-20447, que soluciona un problema que podía provocar problemas de caché, especialmente cuando las regiones estaban divididas.
- MXnet
  - Se han agregado bibliotecas OpenCV.
- Spark
  - Cuando Spark escribe archivos Parquet en una ubicación de Amazon S3 mediante EMRFS, el FileOutputCommitter algoritmo se ha actualizado para usar la versión 2 en lugar de la versión 1. Esto reduce el número de cambios de nombre, lo que mejora el rendimiento de la aplicación. Este cambio no afecta a:
    - Aplicaciones distintas de Spark.
    - Aplicaciones que escriben en otros sistemas de archivos, como HDFS (que aún utilizan la versión 1 de). FileOutputCommitter
    - Aplicaciones que utilizan otros formatos de salida, como texto o csv, que ya utilizan la escritura directa de EMRFS.

Problemas conocidos

- JupyterHub
  - No se admite el uso de clasificaciones de configuración para configurar JupyterHub cuadernos de Jupyter individuales al crear un clúster. Edite manualmente los archivos jupyterhub_config.py

y jupyter_notebook_config.py para cada usuario. Para obtener más información, consulte Configurando JupyterHub.

- JupyterHub no se inicia en los clústeres de una subred privada y se produce un error con el mensaje. `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` Esto es provocado por un error en el script que genera los certificados autofirmados. Utilice la siguiente solución alternativa para generar certificados autofirmados. Todos los comandos se ejecutan mientras está conectado al nodo principal.

  1. Copie el script de generación de certificados del contenedor al nodo principal:

     ```
     sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
     ```

  2. Utilice un editor de texto para cambiar la línea 23 y cambiar el nombre de host público por un nombre de host local, como se muestra a continuación:

     ```
     local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
     ```

  3. Ejecute el script para generar certificados autofirmados:

     ```
     sudo bash ./gen_self_signed_cert.sh
     ```

  4. Mueva los archivos de certificados que genera el script al directorio `/etc/jupyter/conf/`:

     ```
     sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
     ```

  Puedes revisar `tail` el `jupyter.log` archivo para comprobar que se ha JupyterHub reiniciado y que devuelve un código de respuesta de 200. Por ejemplo:

  ```
  tail -f /var/log/jupyter/jupyter.log
  ```

  Este debe devolver una respuesta similar a la siguiente:

  ```
  # [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
   https://:9443/
  # 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
  ```

- Una vez que se reinicie el nodo principal o se reinicie el controlador de instancias, no se recopilarán las CloudWatch métricas y la función de escalado automático no estará disponible

en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR. Este problema se ha corregido en la versión 5.17.0 de Amazon EMR.

Versiones de los componentes de la versión 5.14.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.23.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.8.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.4.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.2	HBase cliente de línea de comandos.
hbase-rest-server	1.4.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-2	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.2-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
jupyterhub	0.8.1	Servidor multiusuario para blocs de notas Jupyter
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.1.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
opencv	3.4.0	Biblioteca de visión artificial de código abierto.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.194	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.194	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.0	Clientes de línea de comando de Spark.

Componente	Versión	Descripción
spark-history-server	2.3.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.7	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.14.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.14.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
container-log4j	Cambiar los valores en el archivo container-log4j.properties de Hadoop YARN.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
jupyter-notebook-conf	Cambiar los valores en el archivo jupyter_n otebook_config.py de Jupyter Notebook.
jupyter-hub-conf	Cambie los valores en JupyterHubs el archivo jupyterhub_config.py.
jupyter-sparkmagic-conf	Cambiar los valores en el archivo config.json de Sparkmagic.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.13.1

Versiones de las aplicaciones de la versión 5.13.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
AWS SDK for Java	1.11.297	1.11.297	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.0	1.4.0

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.188	0.188
Spark	2.3.0	2.3.0	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.13.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.13.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.22.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.2	HBase cliente de línea de comandos.
hbase-rest-server	1.4.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.2-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.0.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.194	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.194	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.0	Clientes de línea de comando de Spark.
spark-history-server	2.3.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.13.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.13.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.

Clasificaciones	Descripción
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.

Clasificaciones	Descripción
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.13.0

Versiones de las aplicaciones de la versión 5.13.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
AWS SDK for Java	1.11.297	1.11.267	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.188	0.188	0.188
Spark	2.3.0	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.13.0

Las siguientes notas de la versión incluyen información sobre la versión 5.13.0 de Amazon EMR. Los cambios son respecto a la versión 5.12.0.

Actualizaciones

- Se ha actualizado Spark a la versión 2.3.0
- Actualizado HBase a 1.4.2
- Se ha actualizado Presto a la versión 0.194
- Actualizado a la versión AWS SDK para Java 1.11.297

Cambios, mejoras y problemas resueltos

- Hive
  - Se ha adaptado HIVE-15436. Se ha mejorado la colmena APIs para que solo devuelva las vistas.

Problemas conocidos

- MXNet actualmente no tiene bibliotecas OpenCV.

Versiones de los componentes de la versión 5.13.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres

veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.10.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.22.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació

Componente	Versión	Descripción
		n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.2	HBase cliente de línea de comandos.
hbase-rest-server	1.4.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.2	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.3.2-amzn-2	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-2	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-2	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-2	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-2	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark

Componente	Versión	Descripción
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.0.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.194	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.194	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.17.0	Cliente de línea de comando de Pig.
r	3.4.1	Proyecto R para análisis estadístico
spark-client	2.3.0	Clientes de línea de comando de Spark.
spark-history-server	2.3.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.3.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.3.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.13.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.13.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.

Clasificaciones	Descripción
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.

Clasificaciones	Descripción
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.12.3

Versiones de las aplicaciones de la versión 5.12.3

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
AWS SDK for Java	1.11.267	1.11.267	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.4.0

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.188
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.12.3

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.12.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.9.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.21.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.0	HBase cliente de línea de comandos.
hbase-rest-server	1.4.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.2-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.0.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.188	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.188	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.12.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.12.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop

Clasificaciones	Descripción
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.

Clasificaciones	Descripción
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.12.2

Versiones de las aplicaciones de la versión 5.12.2

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
AWS SDK for Java	1.11.267	1.11.267	1.11.267	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.12.2

Las siguientes notas de la versión incluyen información sobre la versión 5.12.2 de Amazon EMR. Los cambios son respecto a la versión 5.12.1.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- Esta versión aborda una posible vulnerabilidad de seguridad.

Versiones de los componentes de la versión 5.12.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.9.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.21.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.8.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.4.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.0	HBase cliente de línea de comandos.
hbase-rest-server	1.4.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-1	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.2-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.0.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.3.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.188	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.188	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.12.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.12.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.

Clasificaciones	Descripción
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.12.1

Versiones de las aplicaciones de la versión 5.12.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
AWS SDK for Java	1.11.267	1.11.267	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	0.12.0	0.12.0

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.12.1

Las siguientes notas de la versión incluyen información sobre la versión 5.12.1 de Amazon EMR. Los cambios son respecto a la versión 5.12.0.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versiones de los componentes de la versión 5.12.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.9.0	Aplicación de copia distribuida optimizada para Amazon S3.

Componente	Versión	Descripción
emrfs	2.21.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.8.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.8.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.4.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.0	HBase cliente de línea de comandos.
hbase-rest-server	1.4.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-1	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.2-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	1.0.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.3.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.188	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.188	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.12.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.12.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.

Clasificaciones	Descripción
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.12.0

Versiones de las aplicaciones de la versión 5.12.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
AWS SDK for Java	1.11.267	1.11.238	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	0.12.0	0.12.0	0.12.0

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.12.0

Las siguientes notas de la versión incluyen información sobre la versión 5.12.0 de Amazon EMR. Los cambios son respecto a la versión 5.11.1.

Actualizaciones

- AWS SDK para Java 1.11.238 ⇒ 1.11.267. Para obtener más información, consulta el registro de cambios de AWS SDK for Java GitHub.
- Hadoop 2.7.3 ⇒ 2.8.3. Para más información, consulte Versiones de Apache Hadoop.
- Flink 1.3.2 ⇒ 1.4.0. Para más información, consulte Apache Flink 1.4.0 Release Announcement.
- HBase 1.3.1 ⇒ 1.4.0. Para obtener más información, consulta el anuncio de HBase lanzamiento.

- Hue 4.0.1 ⇒ 4.1.0. Para más información, consulte las notas de la versión.
- MxNet 0.12.0 ⇒ 1.0.0. Para obtener más información, consulte el registro de MXNet cambios. GitHub
- Presto 0.187 ⇒ 0.188. Para más información, consulte las notas de la versión.

Cambios, mejoras y problemas resueltos

- Hadoop
  - La propiedad `yarn.resourcemanager.decommissioning.timeout` se ha cambiado a `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Puede utilizar esta propiedad para personalizar la reducción vertical del clúster. Para más información, consulte Reducción vertical de clúster en la Guía de administración de Amazon EMR.
  - La CLI de Hadoop ha agregado la opción `-d` al comando `cp` (copia), que especifica la copia directa. Puede utilizarlo para evitar la creación de un archivo `.COPYING` intermediario, lo que agiliza la copia de datos entre Amazon S3. Para más información, consulte HADOOP-12384.
- Pig
  - Se ha agregado la clasificación de configuración `pig-env`, que simplifica la configuración de las propiedades del entorno de Pig. Para obtener más información, consulte Configuración de aplicaciones.
- Presto
  - Se ha agregado la clasificación de configuración `presto-connector-redshift`, que se puede utilizar para configurar valores en el archivo de configuración `redshift.properties` de Presto. Para más información, consulte Redshift Connector en la documentación de Presto y Configuración de aplicaciones.
  - Se ha agregado compatibilidad con Presto para EMRFS y es la configuración predeterminada. Las versiones anteriores de Amazon EMR utilizaban PrestoS3FileSystem, que era la única opción. Para obtener más información, consulte Configuración de EMRFS y PrestoS3PrestoS3 FileSystem.

    > **Note**
    >
    > Si consulta los datos subyacentes en Amazon S3 con la versión 5.12.0 de Amazon EMR, pueden producirse errores en Presto. Esto se debe a que Presto no recoge los valores de clasificación de configuración de `emrfs-site.xml`. Como solución alternativa, cree

un subdirectorio de `emrfs` en `usr/lib/presto/plugin/hive-hadoop2/` y cree un enlace simbólico en `usr/lib/presto/plugin/hive-hadoop2/emrfs` al archivo `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existente. A continuación, reinicie el proceso presto-server (`sudo presto-server stop` seguido de `sudo presto-server start`).

- Spark
  - SPARK-22036 compatible: la multiplicación a veces devuelve un valor nulo. BigDecimal

Problemas conocidos

- MXNet no incluye bibliotecas OpenCV.
- SparkR no está disponible para los clústeres creados con una AMI personalizada porque R no está instalado de forma predeterminada en los nodos del clúster.

Versiones de los componentes de la versión 5.12.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0.1	SDK de Amazon SageMaker Spark

Componente	Versión	Descripción
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.9.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.21.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.4.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.8.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.8.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.8.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.8.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.8.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.8.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.8.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.8.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.8.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.8.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.4.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.4.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.4.0	HBase cliente de línea de comandos.
hbase-rest-server	1.4.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.4.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-1	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.2-amzn-1	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-1	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-1	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-1	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.1.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	1.0.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.1.85	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.13.0- -1.4 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.13.0- -1,4 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.188	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.188	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.12.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.12.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop

Clasificaciones	Descripción
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.

Clasificaciones	Descripción
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-env	Cambiar los valores en el entorno de Pig.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-redshift	Cambiar los valores en el archivo redshift.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.11.4

Versiones de las aplicaciones de la versión 5.11.4

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.11.4

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.11.4

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.8.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.20.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.7.3-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-6	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	0.12.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.0.176	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.3.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.187	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.187	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.11.4

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.11.4

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.

Clasificaciones	Descripción
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.

Clasificaciones	Descripción
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.11.3

Versiones de las aplicaciones de la versión 5.11.3

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.238

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.11.3

Las siguientes notas de la versión incluyen información sobre la versión 5.11.3 de Amazon EMR. Los cambios son respecto a la versión 5.11.2.

Fecha de lanzamiento inicial: 18 de julio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Versiones de los componentes de la versión 5.11.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de

Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.8.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.20.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-6	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-6	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.2-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	0.12.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.0.176	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.3.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.187	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.187	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.11.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.11.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.

Clasificaciones	Descripción
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.

Clasificaciones	Descripción
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.11.2

Versiones de las aplicaciones de la versión 5.11.2

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
AWS SDK for Java	1.11.238	1.11.238	1.11.238	1.11.221

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.11.2

Las siguientes notas de la versión incluyen información sobre la versión 5.11.2 de Amazon EMR. Los cambios son respecto a la versión 5.11.1.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- Esta versión aborda una posible vulnerabilidad de seguridad.

Versiones de los componentes de la versión 5.11.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de

macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.8.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.20.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-6	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.7.3-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	0.12.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.0.176	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.187	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.187	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.11.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.11.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.

Clasificaciones	Descripción
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.

Clasificaciones	Descripción
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.11.1

Versiones de las aplicaciones de la versión 5.11.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
AWS SDK for Java	1.11.238	1.11.238	1.11.221	1.11.221
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.1	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.1	2.3.1

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.11.1

Las siguientes notas de la versión incluyen información sobre la versión 5.11.1 de Amazon EMR. Los cambios son respecto a la versión 5.8.0 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versiones de los componentes de la versión 5.11.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.8.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.20.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-6	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.2-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	0.12.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.0.176	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.187	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.187	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.11.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.11.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop

Clasificaciones	Descripción
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.

Clasificaciones	Descripción
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.11.0

Versiones de las aplicaciones de la versión 5.11.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
AWS SDK for Java	1.11.238	1.11.221	1.11.221	1.11.183
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.1	2.3.1	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.1	2.3.1	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	-

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.184
Spark	2.2.1	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.11.0

Las siguientes notas de la versión incluyen información sobre la versión 5.11.0 de Amazon EMR. Los cambios son respecto a la versión 5.10.0.

Actualizaciones

- Hive 2.3.2
- Spark 2.2.1
- SDK para Java 1.11.238

Nuevas características

- Spark
  - Se ha agregado la configuración `spark.decommissioning.timeout.threshold`, que mejora el comportamiento de retirada de Spark cuando se utilizan instancias de spot. Para obtener más información, consulte Configuración del comportamiento de retirada de nodos.
  - Se agregó el `aws-sagemaker-spark-sdk` componente a Spark, que instala Amazon SageMaker Spark y las dependencias asociadas para la integración de Spark con Amazon. SageMaker Puedes usar Amazon SageMaker Spark para crear canalizaciones de aprendizaje automático (ML) de Spark mediante Amazon SageMaker Stages. Para obtener más información, consulta el Léame de SageMaker Spark sobre el GitHub uso de Apache Spark con Amazon SageMaker en la Guía para SageMaker desarrolladores de Amazon.

Problemas conocidos

- MXNet no incluye bibliotecas OpenCV.
- Hive 2.3.2 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

  Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

Versiones de los componentes de la versión 5.11.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
aws-sagemaker-spark-sdk	1.0	SDK de Amazon SageMaker Spark
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.8.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.20.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto

Componente	Versión	Descripción
		con el agente de monitoriz ación de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-6	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-6	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-6	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-6	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-6	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.7.3-amzn-6	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-6	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-6	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-6	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-6	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.2-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.2-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.2-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.2-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.2-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mxnet	0.12.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.0.176	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.187	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.187	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.1	Clientes de línea de comando de Spark.
spark-history-server	2.2.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.11.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.11.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.

Clasificaciones	Descripción
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.

Clasificaciones	Descripción
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.10.1

Versiones de las aplicaciones de la versión 5.10.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
AWS SDK for Java	1.11.221	1.11.221	1.11.183	1.11.183
Python	2.7, 3.4	2.7, 3.4	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.1	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.1	2.3.0	2.3.0

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.184	0.184
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.10.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.10.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.7.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.20.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-5	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	0.12.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.0.176	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.187	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.187	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.10.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.10.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop

Clasificaciones	Descripción
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.

Clasificaciones	Descripción
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.10.0

Versiones de las aplicaciones de la versión 5.10.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout,, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TezZeppelin ZooKeeper.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
AWS SDK for Java	1.11.221	1.11.183	1.11.183	1.11.160
Python	2.7, 3.4	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	-
MXNet	0.12.0	-	-	-

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.16,0
Presto	0.187	0.184	0.184	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.10.0

Las siguientes notas de la versión incluyen información sobre la versión 5.10.0 de Amazon EMR. Los cambios son respecto a la versión 5.9.0 de Amazon EMR.

Actualizaciones

- AWS SDK para Java 1.11.221
- Hive 2.3.1
- Presto 0.187

Nuevas características

- Se ha agregado compatibilidad para autenticación Kerberos. Para más información, consulte Uso de la autenticación de Kerberos en la Guía de administración de Amazon EMR
- Se ha agregado compatibilidad con los roles de IAM para EMRFS. Para más información, consulte Configurar roles de IAM de solicitudes de EMRFS para Amazon S3 en la Guía de administración de Amazon EMR
- Se ha agregado un soporte para los tipos de instancias P2 y P3 basados en GPU. Para obtener más información, consulte Amazon EC2 P2 Instances e Amazon EC2 P3 Instances. El controlador NVIDIA 384,81 y el controlador CUDA 9.0.176 están instalados en estos tipos de instancias de forma predeterminada.
- Se agregó compatibilidad con Apache MXNet.

Cambios, mejoras y problemas resueltos

- Presto
  - Se ha añadido compatibilidad con el uso del catálogo de datos de AWS Glue como metastore predeterminado de Hive. Para obtener más información, consulte Uso de Presto con el catálogo de datos de AWS Glue.
  - Se ha agregado compatibilidad para las funciones geoespaciales.
  - Se ha agregado compatibilidad con las uniones de derrame en el disco.
  - Se ha agregado compatibilidad con el conector de Redshift.
- Spark
  - Se ha adaptado SPARK-20640, lo que hace que el tiempo de espera de rpc y los reintentos de los valores de registro de reorganización sean configurables mediante las propiedades `spark.shuffle.registration.timeout` y `spark.shuffle.registration.maxAttempts`.
  - El SPARK-21549 es compatible, lo que corrige un error que se producía al grabar de forma personalizada en ubicaciones que no son HDFS. OutputFormat
- Se ha adaptado Hadoop-13270
- Se han eliminado las bibliotecas Numpy, Scipy y Matplotlib de la AMI base de Amazon EMR. Si la aplicación requiere estas bibliotecas, están disponibles en el repositorio de aplicaciones, por lo que puede utilizar una acción de arranque para instalarlas en todos los nodos mediante `yum install`.

- La AMI base de Amazon EMR ya no incluye paquetes RPM de aplicaciones, por lo que los paquetes RPM ya no están presentes en los nodos del clúster. Custom AMIs y la AMI base de Amazon EMR ahora hacen referencia al repositorio de paquetes RPM de Amazon S3.
- Gracias a la introducción de la facturación por segundo en Amazon EC2, el comportamiento de reducción predeterminado ahora es Finalizar al finalizar la tarea en lugar de Finalizar a la hora de la instancia. Para obtener más información, consulte Reducción de escala del clúster.

Problemas conocidos

- MXNet no incluye bibliotecas OpenCV.
- Hive 2.3.1 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

  Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

Versiones de los componentes de la versión 5.10.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.5.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.7.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.20.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-5	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-5	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-5	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-5	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-5	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-5	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-5	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-5	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-5	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-5	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mxnet	0.12.0	Una biblioteca flexible, escalable y eficiente para el aprendizaje profundo.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
nvidia-cuda	9.0.176	Controladores Nvidia y conjunto de herramientas Cuda
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.187	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.187	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.3	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.10.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.10.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop

Clasificaciones	Descripción
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.

Clasificaciones	Descripción
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.9.1

Versiones de las aplicaciones de la versión 5.9.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia,, HBase, HCatalog, Hadoop, Hive, Hue,, Livy, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
AWS SDK for Java	1.11.183	1.11.183	1.11.160	1.11.160
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	-	-
MXNet	-	-	-	-

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.16,0	0,16.0
Presto	0.184	0.184	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.9.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.9.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.4.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.7.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.19.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-4	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.0-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.0-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4,11,0- -1,3 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.184	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.184	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.9.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.9.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.

Clasificaciones	Descripción
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.

Clasificaciones	Descripción
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.

Clasificaciones	Descripción
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.9.0

Versiones de las aplicaciones de la versión 5.9.0

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia,, HBase, HCatalog, Hadoop, Hive, Hue,, Livy, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
AWS SDK for Java	1.11.183	1.11.160	1.11.160	1.11.160
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.16,0	0,16,0	0,16,0
Presto	0.184	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.9.0

Las siguientes notas de la versión incluyen información sobre la versión 5.9.0 de Amazon EMR. Los cambios son respecto a la versión 5.8.0 de Amazon EMR.

Fecha de lanzamiento: 5 de octubre de 2017

Última actualización de características: 12 de octubre de 2017

Actualizaciones

- AWS SDK para Java versión 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

Nuevas características

- Se ha añadido compatibilidad con Livy (versión 0.4.0-incubating). Para obtener más información, consulte Apache Livy.
- Se ha añadido compatibilidad con Hue Notebook para Spark.
- Se agregó soporte para las EC2 instancias Amazon de la serie i3 (12 de octubre de 2017).

Cambios, mejoras y problemas resueltos

- Spark

  - Se ha añadido un nuevo conjunto de características que ayudan a asegurarse de que Spark gestiona de un modo más correcto la terminación de los nodos debido a una solicitud de cambio de tamaño manual o a una política de escalado automático. Para obtener más información, consulte Configuración del comportamiento de retirada de nodos.
  - Se utiliza SSL en lugar de 3DES para el cifrado en tránsito del servicio de transferencia de bloques, lo que mejora el rendimiento cuando se utilizan tipos de EC2 instancias de Amazon con AES-NI.
  - Se ha adaptado SPARK-21494.
- Zeppelin
  - Se ha adaptado ZEPPELIN-2377.
- HBase
  - Se agregó el parche HBASE-18533, que permite valores adicionales para la configuración mediante la clasificación de configuración. HBase BucketCache `hbase-site`
- Hue
  - Se agregó compatibilidad con AWS Glue Data Catalog para el editor de consultas Hive en Hue.
  - Los superusuarios de Hue pueden acceder, de forma predeterminada, a todos los archivos para los que tienen autorización de acceso los roles de IAM de Amazon EMR. Los usuarios recién creados no obtienen automáticamente permisos de acceso al navegador de archivos de Amazon S3, por lo que es necesario activar los permisos `filebrowser.s3_access` para su grupo.
- Se ha resuelto un problema que provocaba que no se pudiera acceder a los datos JSON subyacentes creados con AWS Glue Data Catalog.

Problemas conocidos

- El lanzamiento del clúster produce un error cuando se instalan todas las aplicaciones y no se cambia el tamaño del volumen raíz de Amazon EBS predeterminado. Como solución alternativa, utilice el `aws emr create-cluster` comando del AWS CLI y especifique un parámetro mayor`--ebs-root-volume-size`.
- Hive 2.3.0 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

Versiones de los componentes de la versión 5.9.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.4.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.7.0	Aplicación de copia distribuida optimizada para Amazon S3.

Componente	Versión	Descripción
emrfs	2.19.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.2	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-4	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.0-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.0-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.3.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	4.0.1	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
livy-server	0.4.0-incubating	Interfaz de REST para interactuar con Apache Spark
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.184	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.184	Service para ejecutar partes de una consulta.
pig-client	0.17.0	Cliente de línea de comando de Pig.
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.9.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.9.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
livy-conf	Cambiar los valores en el archivo livy.conf de Livy.
livy-env	Cambiar los valores en el entorno de Livy.
livy-log4j	Cambiar la configuración de log4j.properties de Livy.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.

Clasificaciones	Descripción
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.

Clasificaciones	Descripción
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.

Clasificaciones	Descripción
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.8.3

Versiones de las aplicaciones de la versión 5.8.3

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
AWS SDK for Java	1.11.160	1.11.160	1.11.160	1.11.160
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.8.3

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.8.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.4.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.6.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.18.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.0-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.0-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-1	Cliente de línea de comando de Pig.
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.8.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.8.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.

Clasificaciones	Descripción
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.8.2

Versiones de las aplicaciones de la versión 5.8.2

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
AWS SDK for Java	1.11.160	1.11.160	1.11.160	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0,16,0	0,16,0	0,16,0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.1.1

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.8.2

Las siguientes notas de la versión incluyen información sobre la versión 5.8.2 de Amazon EMR. Los cambios son respecto a la versión 5.8.1.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versiones de los componentes de la versión 5.8.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*-

`amzn-EmrVersion`. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.4.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.6.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.18.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.7.3-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.0-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.0-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-1	Cliente de línea de comando de Pig.
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.8.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.8.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.

Clasificaciones	Descripción
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.8.1

Versiones de las aplicaciones de la versión 5.8.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
AWS SDK for Java	1.11.160	1.11.160	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.0	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.1.1	2.1.1

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.8.1

Las siguientes notas de la versión incluyen información sobre la versión 5.8.1 de Amazon EMR. Los cambios son respecto a la versión 5.8.0 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versiones de los componentes de la versión 5.8.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.4.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.6.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.18.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.3.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.3.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.0-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.0-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-1	Cliente de línea de comando de Pig.
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.

Componente	Versión	Descripción
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.8.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.8.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.

Clasificaciones	Descripción
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.8.0

Versiones de las aplicaciones de la versión 5.8.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
AWS SDK for Java	1.11.160	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
Delta	-	-	-	-
Flink	1.3.1	1.3.0	1.3.0	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.0
HCatalog	2.3.0	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.1.1	2.1.1	2.1.1

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.8.0

Las siguientes notas de la versión incluyen información sobre la versión 5.8.0 de Amazon EMR. Los cambios son respecto a la versión 5.7.0 de Amazon EMR.

Fecha de lanzamiento inicial: 10 de agosto de 2017

Última actualización de características: 25 de septiembre de 2017

Actualizaciones

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0. Para obtener más información, consulte Release Notes en el sitio de Apache Hive.
- Spark 2.2.0. Para obtener más información, consulte Release Notes en el sitio de Apache Spark.

Nuevas características

- Se ha agregado la posibilidad de visualizar el historial de aplicaciones (25 de septiembre de 2017). Para más información, consulte Visualización del historial de aplicaciones en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Integración con AWS Glue Data Catalog
  - Se ha añadido la posibilidad de que Hive y Spark SQL utilicen AWS Glue Data Catalog como almacén de metadatos de Hive. Para obtener más información, consulte Se ha agregado compatibilidad para utilizar el Catálogo de datos de AWS Glue como metaalmacén predeterminado de Hive y Usa el catálogo de AWS Glue Data Catalog con Spark en Amazon EMR.
- Se ha añadido el Application history (Historial de aplicaciones) a los detalles del clúster, lo que le permite ver datos históricos de las aplicaciones de YARN y detalles adicionales de las aplicaciones de Spark. Para obtener más información, consulte Ver el historial de aplicaciones en la Guía de administración de Amazon EMR.
- Oozie
  - Se ha adaptado OOZIE-2748.
- Hue
  - Se ha adaptado HUE-5859
- HBase
  - Se agregó un parche para mostrar la hora de inicio del servidor HBase maestro mediante el uso de Java Management Extensions (JMX). `getMasterInitializedTime`
  - Se ha añadido un parche que mejora la hora de inicio del clúster.

Problemas conocidos

- El lanzamiento del clúster produce un error cuando se instalan todas las aplicaciones y no se cambia el tamaño del volumen raíz de Amazon EBS predeterminado. Como solución alternativa, utilice el `aws emr create-cluster` comando del AWS CLI y especifique un parámetro mayor`--ebs-root-volume-size`.
- Hive 2.3.0 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

- Spark: cuando se utiliza Spark, hay un problema de fuga de un controlador de archivos con el daemon apppusher que puede aparecer con un trabajo de Spark de ejecución prolongada después de varias horas o días. Para solucionar el problema, conéctese al nodo principal y escriba `sudo /etc/init.d/apppusher stop`. Esto detiene el daemon apppusher, que Amazon EMR reiniciará automáticamente.
- Historial de aplicaciones
  - Los datos históricos de los ejecutores de Spark inactivos no están disponibles.
  - El historial de aplicaciones no está disponible para clústeres que utilizan una configuración de seguridad para habilitar el cifrado en tránsito.

Versiones de los componentes de la versión 5.8.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.4.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.4.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.4.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.6.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.18.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.3.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.3.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.3.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.3.0-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.3.0-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.3.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-1	Cliente de línea de comando de Pig.
spark-client	2.2.0	Clientes de línea de comando de Spark.
spark-history-server	2.2.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.2.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.2.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.8.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.8.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.

Clasificaciones	Descripción
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.7.1

Versiones de las aplicaciones de la versión 5.7.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
Delta	-	-	-	-
Flink	1.3.0	1.3.0	1.2.1	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.1

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.7.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.7.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.18.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.

Componente	Versión	Descripción
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.1	Clientes de línea de comando de Spark.
spark-history-server	2.1.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.

Componente	Versión	Descripción
spark-on-yarn	2.1.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.7.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.7.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.

Clasificaciones	Descripción
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.7.0

Versiones de las aplicaciones de la versión 5.7.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
Delta	-	-	-	-
Flink	1.3.0	1.2.1	1.2.1	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.0

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.7.0

Las siguientes notas de la versión incluyen información sobre la versión 5.7.0 de Amazon EMR. Los cambios son respecto a la versión 5.6.0 de Amazon EMR.

Fecha de lanzamiento: 13 de julio de 2017

Actualizaciones

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

Nuevas características

- Se ha agregado la capacidad de especificar una AMI personalizada de Amazon Linux al crear un clúster. Para obtener más información, consulte Uso de una AMI personalizada.

Cambios, mejoras y problemas resueltos

- HBase
  - Se agregó la capacidad de configurar clústeres de HBase lectura-réplica. Consulte Uso de un clúster de réplicas de lectura.

  - Varias correcciones de errores y mejoras
- Presto: se ha agregado la capacidad de configurar `node.properties`.
- YARN: se ha agregado la capacidad de configurar `container-log4j.properties`
- Sqoop: se ha adaptado SQOOP-2880, que introduce un argumento que le permite definir el directorio temporal Sqoop.

Versiones de los componentes de la versión 5.7.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.18.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.3.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.11.0- -1.3 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4.11.0- -1.3 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.1	Clientes de línea de comando de Spark.
spark-history-server	2.1.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.7.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.7.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.6.1

Versiones de las aplicaciones de la versión 5.6.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppelin y. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.1	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.6.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.6.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.17.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.1	Clientes de línea de comando de Spark.
spark-history-server	2.1.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.1.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.6.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.6.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.

Clasificaciones	Descripción
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.prope rties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.pro perties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.6.0

Versiones de las aplicaciones de la versión 5.6.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
Delta	-	-	-	-
Flink	1.2.1	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.0	2.1.0	2.1.0

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.6.0

Las siguientes notas de la versión incluyen información sobre la versión 5.6.0 de Amazon EMR. Los cambios son respecto a la versión 5.5.0 de Amazon EMR.

Fecha de lanzamiento: 5 de junio de 2017

Actualizaciones

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Esta es la primera versión de Mahout que admite Spark 2.x en la versión 5.0 y posteriores de Amazon EMR.
- Spark 2.1.1

Cambios, mejoras y problemas resueltos

- Presto
  - Se agregó la capacidad de habilitar la comunicación SSL/TLS segura entre los nodos de Presto al habilitar el cifrado en tránsito mediante una configuración de seguridad. Para obtener más información, consulte Cifrado de datos en tránsito.

- Se ha adaptado Presto 7661, que agrega la opción `VERBOSE` a la instrucción `EXPLAIN ANALYZE` para informar sobre estadísticas de bajo nivel más detalladas sobre un plan de consulta.

Versiones de los componentes de la versión 5.6.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.

Componente	Versión	Descripción
emrfs	2.17.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.1	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.3.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.1	HBase cliente de línea de comandos.
hbase-rest-server	1.3.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.1.1-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.13.0	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.1	Clientes de línea de comando de Spark.
spark-history-server	2.1.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.6.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.6.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.

Clasificaciones	Descripción
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.

Clasificaciones	Descripción
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-env	Cambiar valores en el archivo presto-env.sh de Presto.
presto-node	Cambiar valores en el archivo node.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.

Clasificaciones	Descripción
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.5.4

Versiones de las aplicaciones de la versión 5.5.4

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.5.4

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.5.4

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.16.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.3.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.0	HBase cliente de línea de comandos.
hbase-rest-server	1.3.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.5.4

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.5.4

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.

Clasificaciones	Descripción
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.5.3

Versiones de las aplicaciones de la versión 5.5.3

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

Notas de la versión 5.5.3

Las siguientes notas de la versión incluyen información sobre la versión 5.5.3 de Amazon EMR. Los cambios son respecto a la versión 5.5.2.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- Esta versión aborda una posible vulnerabilidad de seguridad.

Versiones de los componentes de la versión 5.5.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.16.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.3.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.0	HBase cliente de línea de comandos.
hbase-rest-server	1.3.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.5.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.5.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.

Clasificaciones	Descripción
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.5.2

Versiones de las aplicaciones de la versión 5.5.2

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.11.0
Iceberg	-	-	-	-

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.170	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.9

Notas de la versión 5.5.2

Las siguientes notas de la versión incluyen información sobre la versión 5.5.2 de Amazon EMR. Los cambios son respecto a la versión 5.5.1.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versiones de los componentes de la versión 5.5.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.16.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.3.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.0	HBase cliente de línea de comandos.
hbase-rest-server	1.3.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.5.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.5.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.

Clasificaciones	Descripción
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.5.1

Versiones de las aplicaciones de la versión 5.5.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.11.0	3.11.0
Iceberg	-	-	-	-

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.170	0.166	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.0	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.9	3.4.9

Notas de la versión 5.5.1

Las siguientes notas de la versión incluyen información sobre la versión 5.5.1 de Amazon EMR. Los cambios son respecto a la versión 5.5.0 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versiones de los componentes de la versión 5.5.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.16.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.3.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.0	HBase cliente de línea de comandos.
hbase-rest-server	1.3.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente

Componente	Versión	Descripción
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.

Componente	Versión	Descripción
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.5.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.5.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.

Clasificaciones	Descripción
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-me trics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-l ogging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.

Clasificaciones	Descripción
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.5.0

Versiones de las aplicaciones de la versión 5.5.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.170	0.166	0.166	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.0	0.7.0	0.6.2
ZooKeeper	3.4.10	3.4.9	3.4.9	3.4.9

Notas de la versión 5.5.0

Las siguientes notas de la versión incluyen información sobre la versión 5.5.0 de Amazon EMR. Los cambios son respecto a la versión 5.4.0 de Amazon EMR.

Fecha de lanzamiento: 26 de abril de 2017

Actualizaciones

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1
- ZooKeeper 3.4,10

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha corregido un parche de Spark con una versión anterior (SPARK-20115) DAGScheduler que permitía volver a calcular todos los bloques de reproducción aleatoria perdidos cuando el servicio de reproducción aleatoria externo no estaba disponible en la versión 2.1.0 de Spark, incluida en esta versión.
- Flink
  - Flink ahora se compila con Scala 2.11. Si utiliza las bibliotecas y la API de Scala, le recomendamos que utilice Scala 2.11 en sus proyectos.
  - Se ha resuelto un problema donde los valores predeterminados de `HADOOP_CONF_DIR` y `YARN_CONF_DIR` no se definían correctamente, por tanto no funcionaba `start-scala-shell.sh`. También se ha añadido la capacidad de definir estos valores utilizando `env.hadoop.conf.dir` y `env.yarn.conf.dir` en `/etc/flink/conf/flink-conf.yaml` o la clasificación de configuración `flink-conf`.
  - Se ha introducido un nuevo comando específico de EMR, `flink-scala-shell` como encapsulador para `start-scala-shell.sh`. Recomendamos utilizar este comando en lugar de `start-scala-shell`. El nuevo comando simplifica la ejecución. Por ejemplo, `flink-scala-shell -n 2` inicia un shell Scala de Flink con un paralelismo de tareas de 2.
  - Se ha introducido un nuevo comando específico de EMR, `flink-yarn-session` como encapsulador para `yarn-session.sh`. Recomendamos utilizar este comando en lugar de `yarn-session`. El nuevo comando simplifica la ejecución. Por ejemplo, `flink-yarn-session -d -n 2` inicia una sesión de Flink de ejecución prolongada en un estado desasociado con dos administradores de tareas.
  - Se ha corregido (FLINK-6125) Commons httpclient is not shaded anymore in Flink 1.2.
- Presto

  - Se ha agregado compatibilidad para autenticación LDAP. El uso de LDAP con Presto en Amazon EMR requiere habilitar el acceso HTTPS para el coordinador de Presto (`http-server.https.enabled=true` en `config.properties`). Para conocer los detalles de configuración, consulte LDAP Authentication en la documentación de Presto.
  - Se agregó compatibilidad con `SHOW GRANTS`.
- AMI base de Linux de Amazon EMR
  - Las versiones de Amazon EMR ahora se basan en Amazon Linux 2017.03. Para obtener más información, consulte las notas de la versión de Amazon Linux AMI 2017.03.
  - Se ha eliminado Python 2.6 de la imagen de Linux de base en Amazon EMR. Python 2.7 y 3.4 se instalan de forma predeterminada. Puede instalar Python 2.6 manualmente si es necesario.

Versiones de los componentes de la versión 5.5.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.5.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.16.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.0	HBase cliente de línea de comandos.
hbase-rest-server	1.3.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.12.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.170	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.170	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.10	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.10	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.5.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.5.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.4.1

Versiones de las aplicaciones de la versión 5.4.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.2.3	1.2.3

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.166	0.166	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.7.0	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 5.4.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.4.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.15.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.3.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.0	HBase cliente de línea de comandos.
hbase-rest-server	1.3.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.1.1-amzn-0	Hive-hbase client.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.166	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.166	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.0	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.4.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.4.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.4.0

Versiones de las aplicaciones de la versión 5.4.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.1.4	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.2.3	1.2.3	1.2.3

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.166	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 5.4.0

Las siguientes notas de la versión incluyen información sobre la versión 5.4.0 de Amazon EMR. Los cambios son respecto a la versión 5.3.0 de Amazon EMR.

Fecha de lanzamiento: 8 de marzo de 2017

Actualizaciones

- Actualizado a Flink 1.2.0
- Actualizado a 1.3.0 HBase
- Actualizado a Phoenix 4.9.0

> **Note**
>
> Si actualiza desde una versión anterior de Amazon EMR a una versión 5.4.0 o posterior de Amazon EMR y utiliza la indexación secundaria, actualice los índices locales como se describe en la documentación de Apache Phoenix. Amazon EMR elimina las configuraciones requeridas de la clasificación de `hbase-site`, pero los índices se tienen que volver a rellenar. Se admite la actualización de índices en línea y fuera de línea. Las actualizaciones en línea son la opción predeterminada, lo que significa que los índices se rellenan al inicializar desde los clientes de Phoenix versión 4.8.0 o superior. Para especificar actualizaciones fuera de línea, defina la configuración `phoenix.client.localIndexUpgrade` como “falso” en la clasificación `phoenix-site` y, a continuación, SSH en el nodo principal para ejecutar `psql [zookeeper] -1`.

- Actualizado a Presto 0.166
- Actualizado a Zeppelin 0.7.0

Cambios y mejoras

- Se ha añadido compatibilidad para instancias r4. Consulte Tipos de EC2 instancias de Amazon.

Versiones de los componentes de la versión 5.4.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.3.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.

Componente	Versión	Descripción
emrfs	2.15.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.2.0	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.3.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.3.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.3.0	HBase cliente de línea de comandos.
hbase-rest-server	1.3.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.3.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-hbase	2.1.1-amzn-0	Hive-hbase client.

Componente	Versión	Descripción
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server2	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.9.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4,9,0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.166	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.166	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.7.0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.4.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.4.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.

Clasificaciones	Descripción
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.

Clasificaciones	Descripción
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.3.2

Versiones de las aplicaciones de la versión 5.3.2

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.4	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 5.3.2

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.3.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.14.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
flink-client	1.1.4	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.3.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.3.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.

Clasificaciones	Descripción
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.

Clasificaciones	Descripción
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.

Clasificaciones	Descripción
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.3.1

Versiones de las aplicaciones de la versión 5.3.1

Esta versión incluye las siguientes aplicaciones: Flink,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 5.3.1

Las siguientes notas de la versión incluyen información sobre la versión 5.3.1 de Amazon EMR. Los cambios son respecto a la versión 5.3.0 de Amazon EMR.

Fecha de lanzamiento: 7 de febrero de 2017

Cambios menores para agregar portabilidad con versiones anteriores de las revisiones de Zeppelin y actualizar la AMI predeterminada de Amazon EMR.

Versiones de los componentes de la versión 5.3.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.14.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.4	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.3.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.3.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.

Clasificaciones	Descripción
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.3.0

Versiones de las aplicaciones de la versión 5.3.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
HCatalog	2.1.1	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 5.3.0

Las siguientes notas de la versión incluyen información sobre la versión 5.3.0 de Amazon EMR. Los cambios son respecto a la versión 5.2.1 de Amazon EMR.

Fecha de lanzamiento: 26 de enero de 2017

Actualizaciones

- Actualizado a Hive 2.1.1
- Actualizado a Hue 3.11.0
- Actualizado a Spark 2.1.0
- Actualizado a Oozie 4.3.0
- Actualizado a Flink 1.1.4

Cambios y mejoras

- Se ha añadido un parche a Hue que le permite utilizar el ajuste `interpreters_shown_on_wheel` para configurar qué intérpretes se muestran primero en la rueda de selección de bloc de notas, independientemente de su orden en el archivo `hue.ini`.
- Se ha agregado la clasificación de configuración `hive-parquet-logging`, que se puede utilizar para configurar valores en el archivo `parquet-logging.properties` de Hive.

Versiones de los componentes de la versión 5.3.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente,

estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.14.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.4	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustad o para aplicaciones del

Componente	Versión	Descripción
		ecosistema de Hadoop junto con el agente de monitoriz ación de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.1-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.1-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.1-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.1-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.1-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.1-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	3.11.0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.3.0	Cliente de línea de comando de Oozie.
oozie-server	4.3.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
spark-client	2.1.0	Clientes de línea de comando de Spark.
spark-history-server	2.1.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.1.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.1.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.3.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.3.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop

Clasificaciones	Descripción
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.

Clasificaciones	Descripción
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-parquet-logging	Cambiar los valores en el archivo parquet-logging.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.

Clasificaciones	Descripción
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.

Clasificaciones	Descripción
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.2.3

Versiones de las aplicaciones de la versión 5.2.3

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.157.1	0.157.1	0.157.1	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.2

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.8

Notas de la versión 5.2.3

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.2.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.13.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.3	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.

Componente	Versión	Descripción
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.

Componente	Versión	Descripción
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.2	Clientes de línea de comando de Spark.
spark-history-server	2.0.2	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.2	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.2.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.2.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.2.2

Versiones de las aplicaciones de la versión 5.2.2

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.157.1	0.157.1	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.8	3.4.8

Notas de la versión 5.2.2

Las siguientes notas de la versión incluyen información sobre la versión 5.2.2 de Amazon EMR. Los cambios son respecto a la versión 5.2.1 de Amazon EMR.

Fecha de lanzamiento: 2 de mayo de 2017

Problemas conocidos resueltos de las versiones anteriores

- Se actualizó el SPARK-194459, que solucionó un problema por el que la lectura de una tabla ORC con columnas podía fallar. char/varchar

Versiones de los componentes de la versión 5.2.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.13.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.3	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.2	Clientes de línea de comando de Spark.
spark-history-server	2.0.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.2	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.2.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.2.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.2.1

Versiones de las aplicaciones de la versión 5.2.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.157.1	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.8	3.4.8	3.4.8

Notas de la versión 5.2.1

Las siguientes notas de la versión incluyen información sobre la versión 5.2.1 de Amazon EMR. Los cambios son respecto a la versión 5.2.0 de Amazon EMR.

Fecha de lanzamiento: 29 de diciembre de 2016

Actualizaciones

- Se ha actualizado Presto a la versión 0.157.1. Para obtener más información, consulte las Presto Release Notes en la documentación de Presto.
- Se ha actualizado ZooKeeper a la versión 3.4.9. Para obtener más información, consulte las notas ZooKeeper de la versión en la documentación de Apache ZooKeeper .

Cambios y mejoras

- Se agregó compatibilidad con el tipo de EC2 instancia Amazon m4.16xlarge en Amazon EMR versión 4.8.3 y posteriores, excepto 5.0.0, 5.0.3 y 5.2.0.
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.
- La ubicación de las rutas de configuración de Flink y YARN ahora se define de forma predeterminada en `/etc/default/flink` y no es necesario definir las variables de entorno `FLINK_CONF_DIR` y `HADOOP_CONF_DIR` al ejecutar los scripts de controlador `flink` o `yarn-session.sh` para lanzar los trabajos de Flink.
- Se agregó soporte para FlinkKinesisConsumer la clase.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema en Hadoop que provocaba que el ReplicationMonitor hilo se quedara atascado durante mucho tiempo debido a una carrera entre la replicación y la eliminación del mismo archivo en un clúster grande.
- Se solucionó un problema por el que ControlledJob #toString fallaba con una excepción de puntero nulo (NPE) cuando el estado del trabajo no se actualizaba correctamente.

Versiones de los componentes de la versión 5.2.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.13.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.3	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.2	Clientes de línea de comando de Spark.
spark-history-server	2.0.2	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.0.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.2.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.2.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.

Clasificaciones	Descripción
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.

Clasificaciones	Descripción
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.

Clasificaciones	Descripción
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.

Clasificaciones	Descripción
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.

Clasificaciones	Descripción
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.2.0

Versiones de las aplicaciones de la versión 5.2.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.152.3	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.1	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 5.2.0

Las siguientes notas de la versión incluyen información sobre la versión 5.2.0 de Amazon EMR. Los cambios son respecto a la versión 5.1.0 de Amazon EMR.

Fecha de lanzamiento: 21 de noviembre de 2016

Cambios y mejoras

- Se agregó el modo de almacenamiento Amazon S3 para HBase.
- Le permite especificar una ubicación de Amazon S3 para el HBase directorio raíz. Para obtener más información, consulte HBase Amazon S3.

Actualizaciones

- Actualizado a Spark 2.0.2

Problemas conocidos resueltos de las versiones anteriores

- Corregido un error con /mnt limitado a 2 TB en tipos de instancias solo para EBS.
- Se ha corregido un error con registros instance-controller y logpusher que generan sus archivos .out correspondientes en lugar de los archivos .log configurados para log4j normales, que rotan cada hora. Los archivos .out no rotan, por lo que esto finalmente rellenaría la partición /emr. Este problema solo afecta a los tipos de instancia de máquina virtual de hardware (HVM).

Versiones de los componentes de la versión 5.2.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.12.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.3	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.152.3	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.152.3	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.2	Clientes de línea de comando de Spark.
spark-history-server	2.0.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.2	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.2.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.2.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase	Configuraciones seleccionadas por Amazon EMR para Apache. HBase
hbase-env	Cambie los valores en su entorno HBase.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .

Clasificaciones	Descripción
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.

Clasificaciones	Descripción
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.

Clasificaciones	Descripción
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.1.1

Versiones de las aplicaciones de la versión 5.1.1

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.2	1.2.2

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16.0	0,16.0	0,16.0
Presto	0.152.3	0.152.3	0.152.3	0.150
Spark	2.0.1	2.0.1	2.0.1	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 5.1.1

Se trata de una versión de parche para añadir la autenticación AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 5.1.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.11.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.3	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.152.3	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.152.3	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.1	Clientes de línea de comando de Spark.
spark-history-server	2.0.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.1.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.1.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-session .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.

Clasificaciones	Descripción
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.1.0

Versiones de las aplicaciones de la versión 5.1.0

Esta versión incluye las siguientes aplicaciones: Flink,,, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppelin y. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.152.3	0.150	0.150
Spark	2.0.1	2.0.1	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 5.1.0

Las siguientes notas de la versión incluyen información sobre la versión 5.1.0 de Amazon EMR. Los cambios son respecto a la versión 5.0.3 de Amazon EMR.

Fecha de lanzamiento: 3 de noviembre de 2016

Cambios y mejoras

- Se ha añadido soporte para Flink 1.1.3.
- Presto se ha añadido como una opción en la sección de bloc de notas de Hue.

Actualizaciones

- Actualizado a 1.2.3 HBase
- Actualizado a Zeppelin 0.6.2

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema con las consultas de Tez en Amazon S3 con archivos de ORC que no tienen un rendimiento tan bueno como en las versiones de Amazon EMR anteriores a la 4.x.

Versiones de los componentes de la versión 5.1.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.11.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
flink-client	1.1.3	Scripts y aplicaciones de cliente de línea de comando de Apache Flink.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.3	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.3	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.3	HBase cliente de línea de comandos.
hbase-rest-server	1.2.3	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.3	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.152.3	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.152.3	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.1	Clientes de línea de comando de Spark.
spark-history-server	2.0.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.2	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.1.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.1.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
flink-conf	Cambiar la configuración de flink-conf.yaml.
flink-log4j	Cambiar la configuración de log4j.properties de Flink.
flink-log4j-yarn-session	Cambie la configuración de Flink log4 j-yarn-se ssion .properties.
flink-log4j-cli	Cambiar la configuración de log4j-cli.properties de Flink.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.

Clasificaciones	Descripción
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.0.3

Versiones de las aplicaciones de la versión 5.0.3

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 5.0.3

Las siguientes notas de la versión incluyen información sobre la versión 5.0.3 de Amazon EMR. Los cambios son respecto a la versión 5.0.0 de Amazon EMR.

Fecha de lanzamiento: 24 de octubre de 2016

Actualizaciones

- Actualizado a Hadoop 2.7.3
- Actualizado a Presto 0.152.3, que incluye soporte para la interfaz web de Presto. Puede acceder a la interfaz web de Presto en el coordinador de Presto utilizando el puerto 8889. Para obtener más información acerca de la interfaz web de Presto, consulte Web Interface en la documentación de Presto.
- Actualizado a Spark 2.0.1
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versiones de los componentes de la versión 5.0.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.10.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.152.3	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.152.3	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.1	Clientes de línea de comando de Spark.
spark-history-server	2.0.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.0.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.0.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.0.2

Versiones de las aplicaciones de la versión 5.0.2

Esta versión incluye las siguientes aplicaciones: Ganglia,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 5.0.2

Versiones de los componentes de la versión 5.0.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.9.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.2-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.7.2-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.150	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.150	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.0	Clientes de línea de comando de Spark.
spark-history-server	2.0.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1-SNAPSHOT	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.0.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.0.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hiv e-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-s ite.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-l og4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log 4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.0.1

Versiones de las aplicaciones de la versión 5.0.1

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 5.0.1

Versiones de los componentes de la versión 5.0.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *CommunityVersion*-amzn-*EmrVersion*. La *EmrVersion* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.9.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los

Componente	Versión	Descripción
		agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.2-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.7.2-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.150	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.150	Service para ejecutar partes de una consulta.
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.0	Clientes de línea de comando de Spark.
spark-history-server	2.0.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.0	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	2.0.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1-SNAPSHOT	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.0.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.0.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.

Clasificaciones	Descripción
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Amazon EMR, versión 5.0.0

Versiones de las aplicaciones de la versión 5.0.0

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie, Phoenix, Pig, Presto,, SparkSqoopTez, Zeppeliny. ZooKeeper

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS SDK for Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0,16,0	0,16,0	0,16,0	0,16,0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 5.0.0

Fecha de lanzamiento: 27 de julio de 2016

Actualizaciones

- Actualizado a Hive 2.1
- Actualizado a Presto 0.150
- Actualizado a Spark 2.0
- Actualizado a Hue 3.10.0
- Actualizado a Pig 0.16.0
- Actualizado a Tez 0.8.4

- Actualizado a Zeppelin 0.6.1

Cambios y mejoras

- Amazon EMR es compatible con las últimas versiones de código abierto de Hive (versión 2.1) y Pig (versión 0.16.0). Si ha utilizado Hive o Pig en Amazon EMR en el pasado, esto podría afectar a algunos casos de uso. Para obtener más información, consulte Hive y Pig.
- El motor de ejecución predeterminado para Hive y Pig es ahora Tez. Para cambiarlo, debería editar los valores adecuados en las clasificaciones de configuración `hive-site` y `pig-properties`, respectivamente.
- Se ha añadido una característica de depuración de paso mejorada, que le permite ver la causa raíz de los errores de paso en caso de que el servicio pueda determinar la causa. Para más información, consulte Depuración escalonada mejorada en la Guía de administración de Amazon EMR.
- Las aplicaciones que anteriormente terminaban por "-Sandbox" ya no tienen ese sufijo. Esto podría interrumpir la automatización, por ejemplo, si utiliza scripts para lanzar clústeres con estas aplicaciones. En la siguiente tabla se muestran los nombres de aplicación la versión 4.7.2 de Amazon EMR en comparación con su versión 5.0.0.

  Cambios de nombre de aplicación

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-Caja de arena	ZooKeeper

- Spark ya está compilado para Scala 2.11.
- Java 8 es ahora el JVM predeterminado. Todas las aplicaciones se ejecutan utilizando el tiempo de ejecución de Java 8. No hay cambios en el destino de código de bytes de ninguna aplicación. La mayoría de las aplicaciones siguen estableciendo como destino Java 7.

- Zeppelin ahora incluye características de autenticación. Para obtener más información, consulte Zeppelin.
- Se ha añadido soporte para configuraciones de seguridad, que le permiten crear y aplicar opciones de cifrado con más facilidad. Para obtener más información, consulte Cifrado de datos.

Versiones de los componentes de la versión 5.0.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.9.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.2-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-3	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.2-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	2.1.0-amzn-0	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	2.1.0-amzn-0	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	2.1.0-amzn-0	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	2.1.0-amzn-0	Cliente de línea de comando de Hive.
hive-metastore-server	2.1.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server	2.1.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.10.0-amzn-0	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.150	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.150	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.16.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	2.0.0	Clientes de línea de comando de Spark.
spark-history-server	2.0.0	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	2.0.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	2.0.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1-SNAPSHOT	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 5.0.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-5.0.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambie los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j2	Cambie los valores en las propiedades HCatalog HCat log4j2.properties de la Web.
hcatalog-webhcat-site	Cambie los valores del archivo webhcat-site.xml de la HCatalog WebHCat.
hive-beeline-log4j2	Cambiar los valores en el archivo beeline-log4j2.properties de Hive.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j2	Cambie los valores en el archivo hive-exec-log4j2.properties de Hive.

Clasificaciones	Descripción
hive-llap-daemon-log4j2	Cambie los valores en el archivo 4j2.properties de Hive. llap-daemon-log
hive-log4j2	Cambiar los valores en el archivo hive-log4 j2.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la aplicación MapReduce .
mapred-site	Cambie los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.

Clasificaciones	Descripción
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-hive-site	Cambiar los valores en el archivo hive-site.xml de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambie los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambie los valores en el ZooKeeper archivo log4j.properties.

Versiones de lanzamiento de Amazon EMR 4.x

Esta sección contiene las versiones de las aplicaciones, las notas de la versión, las versiones de los componentes y las clasificaciones de configuración de cada versión de lanzamiento de Amazon EMR 4.x.

Al lanzar un clúster, puede elegir entre varias versiones de Amazon EMR. Esto le permite probar y utilizar versiones de la aplicación que cumplan con sus requisitos de compatibilidad. El número de versión se especifica con la etiqueta de versión. Las etiquetas de versión tienen el formato `emr-x.x.x` Por ejemplo, `emr-7.9.0`.

Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Para obtener una tabla completa de las versiones de las aplicaciones incluidas en cada versión de Amazon EMR 4.x, consulte Versiones de la aplicación en las versiones 4.x de Amazon EMR.

Temas

- Versiones de la aplicación en las versiones 4.x de Amazon EMR
- Diferencias en la versiones de lanzamiento de Amazon EMR 4.x
- Amazon EMR, versión 4.9.6
- Amazon EMR, versión 4.9.5
- Amazon EMR, versión 4.9.4
- Amazon EMR, versión 4.9.3
- Amazon EMR, versión 4.9.2
- Amazon EMR, versión 4.9.1
- Amazon EMR, versión 4.8.5
- Amazon EMR, versión 4.8.4
- Amazon EMR, versión 4.8.3
- Amazon EMR, versión 4.8.2
- Amazon EMR, versión 4.8.1
- Amazon EMR, versión 4.8.0
- Amazon EMR, versión 4.7.4
- Amazon EMR, versión 4.7.3
- Amazon EMR, versión 4.7.2
- Amazon EMR, versión 4.7.1
- Amazon EMR, versión 4.7.0
- Amazon EMR, versión 4.6.1

- Amazon EMR, versión 4.6.0
- Amazon EMR, versión 4.5.0
- Amazon EMR, versión 4.4.0
- Amazon EMR, versión 4.3.0
- Amazon EMR, versión 4.2.0
- Amazon EMR, versión 4.1.0
- Amazon EMR, versión 4.0.0

Versiones de la aplicación en las versiones 4.x de Amazon EMR

Para ver una tabla completa con las versiones de las aplicaciones disponibles en cada versión de Amazon EMR 4.x, abra Versiones de la aplicación en las versiones 4.x de Amazon EMR en su navegador.

Diferencias en la versiones de lanzamiento de Amazon EMR 4.x

La documentación de las características de Amazon EMR en la Guía de administración de Amazon EMR especifica la versión de Amazon EMR en la que una característica está disponible, así como las diferencias correspondientes entre las características de Amazon EMR que se remontan a la versión 4.0.0.

A partir de la versión 5.0.0 de Amazon EMR, algunas aplicaciones obtuvieron una actualización de versión importante que cambió los detalles de instalación y de su funcionamiento, y otras pasaron de ser aplicaciones de entorno aislado a aplicaciones nativas. Cada tema de esta sección indica las diferencias importantes específicas de la aplicación cuando se utilizan las versiones 4.x de Amazon EMR.

Temas

- Aplicaciones de entorno aislado
- Consideraciones sobre el uso de Hive en Amazon EMR 4.x
- Consideraciones sobre el uso de Pig en Amazon EMR 4.x

Aplicaciones de entorno aislado

Cuando se utilizan las versiones 4.x de Amazon EMR, algunas aplicaciones se consideran aplicaciones de entorno aislado. Las aplicaciones de entorno aislado son versiones preliminares de

la aplicación que se publicaron durante el lanzamiento inicial de Amazon EMR debido a la demanda. Puede usar la consola o la API para que Amazon EMR instale las aplicaciones de entorno aislado de la misma manera que las aplicaciones nativas, pero las aplicaciones de entorno aislado tienen soporte y documentación limitados. AWS CLI Las aplicaciones de entorno aislado se convirtieron en aplicaciones nativas totalmente compatibles en la versión de lanzamiento de Amazon EMR 5.0.0 y posteriores. A continuación, se indican las aplicaciones de entorno aislado en las versiones de lanzamiento de Amazon EMR 4.x:

- Oozie
- Presto
- Sqoop
- Zeppelin
- ZooKeeper

Al instalar las aplicaciones del entorno de pruebas, los nombres de las aplicaciones se indican con el sufijo `-sandbox`. Por ejemplo, para instalar la versión sandbox de, utilice. *`Presto`* `Presto-sandbox` La instalación puede tardar más que para una aplicación totalmente compatible. El número de versión indicado para cada aplicación en esta sección se corresponde con la versión de la comunidad de la aplicación.

Oozie (versiones de entorno aislado)

Oozie está disponible como una aplicación de entorno aislado a partir de la versión 4.1.0 de Amazon EMR.

Los ejemplos de Oozie no se instalan de forma predeterminada con las versiones del entorno de pruebas. Para instalar los ejemplos, conéctese con SSH al nodo principal y ejecute `install-oozie-examples`.

Información sobre la versión de entorno aislado de Oozie

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.9.6	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode,

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
		hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.9.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.9.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.8.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.8.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.8.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.7.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.7.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.6.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.6.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.5.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.4.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Oozie: versión de entorno aislado	Componentes instalados con el entorno aislado de Oozie
emr-4.3.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.1.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Presto (versiones de entorno aislado)

Presto está disponible como una aplicación de entorno aislado a partir de la versión 4.1.0 de Amazon EMR.

Información sobre la versión de entorno aislado de Presto

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.9.6	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client,

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
		hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.9.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.8.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.2	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.1	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.8.0	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.4	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.3	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.7.2	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.1	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.7.0	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.6.1	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.6.0	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.5.0	0.140	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.4.0	0.136	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.3.0	0.130	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.2.0	0.125	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Presto: versión de entorno aislado	Componentes instalados con el entorno aislado de Presto
emr-4.1.0	0.119	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Sqoop (versiones de entorno aislado)

Sqoop está disponible como una aplicación de entorno aislado a partir de la versión 4.4.0 de Amazon EMR.

Información sobre la versión de entorno aislado de Sqoop

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
emr-4.9.6	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
emr-4.9.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
emr-4.8.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
emr-4.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
emr-4.7.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
emr-4.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Sqoop: versión de entorno aislado	Componentes instalados con el entorno aislado de Sqoop
emr-4.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client
emr-4.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client

Zeppelin (versiones de entorno aislado)

Zeppelin está disponible como una aplicación de entorno aislado a partir de la versión 4.1.0 de Amazon EMR.

Información sobre la versión de entorno aislado de Zeppelin

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.9.6	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.9.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.9.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.8.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.8.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.8.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.4	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.7.3	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.2	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.7.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.6.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.6.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.5.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.4.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Zeppelin: versión de entorno aislado	Componentes instalados con el entorno aislado de Zeppelin
emr-4.3.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.2.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.1.0	0.6.0-SNAPSHOT	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

ZooKeeper (versiones sandbox)

ZooKeeper está disponible como una aplicación de entorno aislado a partir de la versión 4.6.0 de Amazon EMR.

ZooKeeper-Información sobre la versión Sandbox

Etiqueta de versión de Amazon EMR	ZooKeeper-Versión Sandbox	Componentes instalados con - Sandbox ZooKeeper
emr-4.9.6	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Etiqueta de versión de Amazon EMR	ZooKeeper-Versión Sandbox	Componentes instalados con - Sandbox ZooKeeper
		cemanager, zookeeper-client, zookeeper-server
emr-4.9.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper-Versión Sandbox	Componentes instalados con - Sandbox ZooKeeper
emr-4.8.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper-Versión Sandbox	Componentes instalados con - Sandbox ZooKeeper
emr-4.8.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper-Versión Sandbox	Componentes instalados con - Sandbox ZooKeeper
emr-4.7.4	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper-Versión Sandbox	Componentes instalados con - Sandbox ZooKeeper
emr-4.7.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.6.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper-Versión Sandbox	Componentes instalados con - Sandbox ZooKeeper
emr-4.6.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-ht tpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Consideraciones sobre el uso de Hive en Amazon EMR 4.x

En esta sección, se abordan las diferencias que deben tenerse en cuenta cuando se utiliza Hive versión 1.0.0 en las versiones 4.x de Amazon EMR, en comparación con Hive 2.x en las versiones 5.x de Amazon EMR.

No se admiten las transacciones ACID

Hive en las versiones 4.x de Amazon EMR no admite las transacciones ACID con datos de Hive almacenados en Amazon S3 cuando se utilizan las versiones 4.x. Si intenta crear una tabla de transacciones en Amazon S3, se produce una excepción.

Lectura y escritura en tablas de Amazon S3

Hive en las versiones 4.x de Amazon EMR puede escribir directamente en Amazon S3 sin utilizar archivos temporales. De esta forma se mejora el rendimiento, pero también impide que se pueda leer y escribir en la misma tabla de Amazon S3 dentro de la misma instrucción de Hive. Una solución consiste en crear y utilizar una tabla temporal en HDFS.

El siguiente ejemplo muestra cómo utilizar varias instrucciones de Hive para actualizar una tabla en Amazon S3. Las instrucciones crean una tabla temporal en HDFS denominada `tmp` con base en una tabla de Amazon S3 denominada `my_s3_table`. A continuación, la tabla de Amazon S3 se actualizará con el contenido de la tabla temporal.

```
CREATE TEMPORARY TABLE tmp LIKE my_s3_table;
INSERT OVERWRITE TABLE tmp SELECT ....;
```

```
INSERT OVERWRITE TABLE my_s3_table SELECT * FROM tmp;
```

Log4j y Log4j 2

Hive en las versiones de lanzamiento de Amazon EMR 4.x utiliza Log4j. A partir de la versión 5.0.0, se utiliza Log4j 2 de forma predeterminada. Estas versiones pueden requerir diferentes configuraciones del registro. Consulte Apache Log4j 2 para obtener más información.

MapReduce es el motor de ejecución por defecto

Las versiones 4.x de Hive on Amazon EMR MapReduce utilizan como motor de ejecución predeterminado. A partir de la versión 5.0.0 de Amazon EMR, se utiliza Tez de forma predeterminada, que proporciona un mejor rendimiento para la mayoría de los flujos de trabajo.

Autorización de Hive

Hive en las versiones de lanzamiento de Amazon EMR 4.x admite Autorización de Hive para HDFS, pero no para EMRFS y Amazon S3. Los clústeres de Amazon EMR se ejecutan con la autorización desactivada de forma predeterminada.

Comportamiento de combinación de archivos de Hive con Amazon S3

Hive en las versiones 4.x de Amazon EMR combina archivos pequeños al final de un trabajo de solo asignación si `hive.merge.mapfiles` es `true`. La combinación solo se activa si el tamaño de salida medio del trabajo es menor que el valor de `hive.merge.smallfiles.avgsize`. Hive en Amazon EMR tiene exactamente el mismo comportamiento si la ruta de salida final se encuentra en HDFS. Sin embargo, si la ruta de salida se encuentra en Amazon S3, se omite el parámetro `hive.merge.smallfiles.avgsize`. En esta situación, la tarea de combinación siempre se activa si `hive.merge.mapfiles` está definido en `true`.

Consideraciones sobre el uso de Pig en Amazon EMR 4.x

La versión 0.14.0 de Pig se instala en los clústeres creados con versiones 4.x de Amazon EMR. Pig se ha actualizado a la versión 0.16.0 en la versión 5.0.0 de Amazon EMR. Las diferencias significativas se explican a continuación.

Motor de ejecución predeterminado diferente

La versión 0.14.0 de Pig en las versiones 4.x de Amazon EMR utiliza MapReduce como motor de ejecución predeterminado. Pig 0.16.0 y versiones posteriores utilizan Apache Tez. Puede establecer

de forma explícita la clasificación de `pig-properties` configuración que `exectype=mapreduce` se va a utilizar. MapReduce

Funciones definidas por el usuario de Dropped Pig () UDFs

Las versiones personalizadas UDFs que estaban disponibles en las versiones 4.x de Pig on Amazon EMR se eliminaron a partir de Pig 0.16.0. La mayoría de ellas UDFs tienen funciones equivalentes que puede utilizar en su lugar. En la siguiente tabla se enumeran las funciones UDFs omitidas y las equivalentes. Para más información, consulte Funciones integradas en el sitio de Apache Pig.

UDF abandonada	Función equivalente
FORMAT_DT(dtformat, date)	GetHour(fecha), GetMinute (fecha), GetMonth (fecha), GetSecond GetWeek (fecha), GetYear (fecha), GetDay (fecha)
EXTRACT(string, pattern)	REGEX_EXTRACT_ALL(string, pattern)
REPLACE(string, pattern, replacement)	REPLACE(string, pattern, replacement)
DATE_TIME()	ToDate()
DURATION(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween (dt, dt2), SecondsBetween (dt, dt2), (dt, dt2), MonthsBetween (dt, dt2), MinutesBetween (dt, dt2) HoursBetween
EXTRACT_DT(format, date)	GetHour(fecha), GetMinute (fecha), GetMonth (fecha), GetSecond (fecha), GetWeek (fecha), GetYear (fecha) GetDay
OFFSET_DT(date, duration)	AddDuration(fecha, duración), SubtractDuration (fecha, duración)
PERIOD(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween (dt, dt2), SecondsBetween (dt, dt2), (dt, dt2), MonthsBetween (dt, dt2), MinutesBetween (dt, dt2) HoursBetween
CAPITALIZE(string)	UCFIRST(string)

UDF abandonada	Función equivalente
CONCAT_WITH()	CONCAT()
INDEX_OF()	INDEXOF()
LAST_INDEX_OF()	LAST_INDEXOF()
SPLIT_ON_REGEX()	STRSPLT()
UNCAPITALIZE()	LCFIRST()

UDFs Se eliminaron los siguientes sin equivalente: FORMAT (), LOCAL_DATE (), LOCAL_TIME (), CENTER (), LEFT_PAD (), REPEAT (), REPLACE_ONCE (), RIGHT_PAD (), STRIP (), STRIP_END (), STRIP_START (), SWAP_CASE ().

Comandos de Grunt suspendidos

Algunos comandos de Grunt se han dejado de utilizar a partir de Pig 0.16.0. En la siguiente tabla, se muestran comandos de Grunt en Pig 0.14.0 y los comandos equivalentes en la versión actual, si procede.

Pig 0.14.0 y los comandos equivalentes de Grunt actuales

Comando de Grunt de Pig 0.14.0	Comando de Grunt de Pig en 0.16.0 y versiones posteriores
cat non-hdfs-path (< >)	fs -cat < non-hdfs-path >;
cd < non-hdfs-path >;	Sin equivalente
ls < non-hdfs-path >;	fs -ls < non-hdfs-path >;
mover < non-hdfs-path > < non-hdfs-path >;	fs -mv < non-hdfs-path > < non-hdfs-path >;
copiar < non-hdfs-path > < non-hdfs-path >;	fs -cp < non-hdfs-path > < non-hdfs-path >;
copyToLocal <local-path>< non-hdfs-path >;	fs - copyToLocal < non-hdfs-path ><local-path>;

Comando de Grunt de Pig 0.14.0	Comando de Grunt de Pig en 0.16.0 y versiones posteriores
copyFromLocal <local-path>< non-hdfs-path >;	fs - copyFromLocal <local-path>< non-hdfs-path >;
mkdir < non-hdfs-path >;	fs -mkdir < >; non-hdfs-path
rm < >; non-hdfs-path	fs -rm -r -SkipTrash < >; non-hdfs-path
non-hdfs-pathrmf < >;	fs -rm -r -SkipTrash < >; non-hdfs-path

Se ha eliminado la capacidad de directorios de inicio que no son de HDFS

Pig 0.14.0 en las versiones 4.x de Amazon EMR tiene dos mecanismos que permiten a los usuarios distintos del usuario `hadoop`, que no tienen directorios de inicio, ejecutar scripts de Pig. El primer mecanismo es una reserva automática que establece el directorio de trabajo inicial en el directorio raíz si el directorio de inicio no existe. El segundo es una propiedad `pig.initial.fs.name` que le permite cambiar el directorio de trabajo inicial.

Estos mecanismos no están disponibles a partir de la versión 5.0.0 de Amazon EMR y los usuarios deben tener un directorio de inicio en HDFS. Esto no se aplica al usuario de `hadoop`, ya que se aprovisiona un directorio de inicio en el lanzamiento. La ejecución de scripts con pasos de Hadoop jar adopta de manera predeterminada el usuario de Hadoop, a menos que se especifique otro usuario mediante `command-runner.jar`.

Amazon EMR, versión 4.9.6

Versiones de las aplicaciones de la versión 4.9.6

Esta versión incluye las siguientes aplicaciones: Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox, Spark, Sqoop-SandboxTezZeppelin-Sandbox, y. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 4.9.6

Esta es una versión de revisión para agregar la autenticación de AWS Signature 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar solicitudes para Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 4.9.6

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres

veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.17.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-9	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-9	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-9	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-9	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-9	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-9	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.9.6

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.9.6

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.9.5

Versiones de las aplicaciones de la versión 4.9.5

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 4.9.5

Las siguientes notas de la versión incluyen información sobre la versión 4.9.5 de Amazon EMR. Los cambios son respecto a la versión 4.9.4.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- HBase
  - Esta versión aborda una posible vulnerabilidad de seguridad.

Versiones de los componentes de la versión 4.9.5

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.

Componente	Versión	Descripción
emrfs	2.17.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-9	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-9	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-9	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-9	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server	1.0.0-amzn-9	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.9.5

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.9.5

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.

Clasificaciones	Descripción
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.9.4

Versiones de las aplicaciones de la versión 4.9.4

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 4.9.4

Las siguientes notas de la versión incluyen información sobre la versión 4.9.4 de Amazon EMR. Los cambios son respecto a la versión 4.9.3.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versiones de los componentes de la versión 4.9.4

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.17.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-9	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-9	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-9	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-9	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-9	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-9	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.9.4

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.9.4

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.9.3

Versiones de las aplicaciones de la versión 4.9.3

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 4.9.3

Las siguientes notas de la versión incluyen información sobre la versión 4.9.3 de Amazon EMR. Los cambios son respecto a la versión 4.9.2 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versiones de los componentes de la versión 4.9.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.17.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-9	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-9	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-9	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-9	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server	1.0.0-amzn-9	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.9.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.9.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.

Clasificaciones	Descripción
presto-connector-localfile	Cambiar los valores en el archivo localfile .properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.p roperties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.pro perties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresq l.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.pr operties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.pro perties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.9.2

Versiones de las aplicaciones de la versión 4.9.2

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR

- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 4.9.2

Las siguientes notas de la versión incluyen información sobre la versión 4.9.2 de Amazon EMR. Los cambios son respecto a la versión 4.9.1 de Amazon EMR.

Fecha de lanzamiento: 13 de julio de 2017

Se han realizado pequeños cambios, correcciones de errores y mejoras en esta versión.

Versiones de los componentes de la versión 4.9.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.3.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.17.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-9	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-9	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-9	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-metastore-server	1.0.0-amzn-9	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-9	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.9.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.9.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.

Clasificaciones	Descripción
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.

Clasificaciones	Descripción
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.9.1

Versiones de las aplicaciones de la versión 4.9.1

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

Notas de la versión 4.9.1

Las siguientes notas de la versión incluyen información sobre la versión 4.9.1 de Amazon EMR. Los cambios son respecto a la versión 4.8.4 de Amazon EMR.

Fecha de lanzamiento: 10 de abril de 2017

Problemas conocidos resueltos de las versiones anteriores

- Se han adaptado HIVE-9976 y HIVE-10106
- Se ha corregido un problema en YARN donde un gran número de nodos (superior a 2000) y contenedores (superiores a 5000) provocan un out-of-memory error, por ejemplo:"`Exception in thread main java.lang.OutOfMemoryError`".

Cambios y mejoras

- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2017.03. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/.
- Se ha eliminado Python 2.6 de la imagen de Linux de base en Amazon EMR. Puede instalar Python 2.6 manualmente si es necesario.

Versiones de los componentes de la versión 4.9.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.3.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.15.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-2	Punto de enlace HTTP para operaciones HDFS.

Componente	Versión	Descripción
hadoop-kms-server	2.7.3-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.

Componente	Versión	Descripción
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-9	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-9	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-9	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-9	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-9	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-9	Servicio para aceptar consultas de Hive como solicitudes web.

Componente	Versión	Descripción
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.

Componente	Versión	Descripción
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.9.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.9.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.

Clasificaciones	Descripción
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.

Clasificaciones	Descripción
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.

Clasificaciones	Descripción
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.

Clasificaciones	Descripción
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.8.5

Versiones de las aplicaciones de la versión 4.8.5

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.152.3
Spark	1.6.3	1.6.3	1.6.3	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.8

Notas de la versión 4.8.5

Esta es una versión de revisión para agregar la autenticación de AWS Signature 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar solicitudes para Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 4.8.5

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.14.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-8	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-8	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-8	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-8	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-metastore-server	1.0.0-amzn-8	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-8	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.8.5

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.8.5

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.

Clasificaciones	Descripción
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.

Clasificaciones	Descripción
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.8.4

Versiones de las aplicaciones de la versión 4.8.4

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.152.3	0.151
Spark	1.6.3	1.6.3	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.8	3.4.8

Notas de la versión 4.8.4

Las siguientes notas de la versión incluyen información sobre la versión 4.8.4 de Amazon EMR. Los cambios son respecto a la versión 4.8.3 de Amazon EMR.

Fecha de lanzamiento: 7 de febrero de 2017

Se han realizado pequeños cambios, correcciones de errores y mejoras en esta versión.

Versiones de los componentes de la versión 4.8.4

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres

veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.14.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el

Componente	Versión	Descripción
		recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-8	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-8	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-8	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-8	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-8	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-8	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.54+	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.25+	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.8.4

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.8.4

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.8.3

Versiones de las aplicaciones de la versión 4.8.3

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.152.3	0.151	0.151
Spark	1.6.3	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.8	3.4.8	3.4.8

Notas de la versión 4.8.3

Las siguientes notas de la versión incluyen información sobre la versión 4.8.3 de Amazon EMR. Los cambios son respecto a la versión 4.8.2 de Amazon EMR.

Fecha de lanzamiento: 29 de diciembre de 2016

Actualizaciones

- Se ha actualizado Presto a la versión 0.157.1. Para obtener más información, consulte las Presto Release Notes en la documentación de Presto.
- Actualizado a Spark 1.6.3. Para obtener más información, consulte las Spark Release Notes en la documentación de Apache Spark.
- Se ha actualizado a ZooKeeper 3.4.9. Para más información, consulte las Spark ZooKeeperRelease Notes en la ZooKeeper documentación de Apache.

Cambios y mejoras

- Se ha agregado compatibilidad con el tipo de instancia EC2 m4.16xlarge de Amazon EMR en la versión 4.8.3 y posteriores de Amazon EMR, excluyendo las versiones 5.0.0, 5.0.3 y 5.2.0.
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema en Hadoop que provocaba que el ReplicationMonitor hilo se quedara atascado durante mucho tiempo debido a una carrera entre la replicación y la eliminación del mismo archivo en un clúster grande.
- Se solucionó un problema por el que ControlledJob #toString fallaba con una excepción de puntero nulo (NPE) cuando el estado del trabajo no se actualizaba correctamente.

Versiones de los componentes de la versión 4.8.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*-

`amzn-EmrVersion`. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.2.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.13.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.

Componente	Versión	Descripción
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.3-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce

Componente	Versión	Descripción
hadoop-yarn-nodemanager	2.7.3-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-1	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-8	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-8	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-8	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-8	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-8	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-8	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- -1,2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- -1,2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.157.1	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.157.1	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.3	Clientes de línea de comando de Spark.
spark-history-server	1.6.3	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.3	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.3	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.9	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.9	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.8.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.8.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.8.2

Versiones de las aplicaciones de la versión 4.8.2

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.152.3	0.151	0.151	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 4.8.2

Las siguientes notas de la versión incluyen información sobre la versión 4.8.2 de Amazon EMR. Los cambios son respecto a la versión 4.8.0 de Amazon EMR.

Fecha de lanzamiento: 24 de octubre de 2016

Actualizaciones

- Actualizado a Hadoop 2.7.3
- Actualizado a Presto 0.152.3, que incluye soporte para la interfaz web de Presto. Puede acceder a la interfaz web de Presto en el coordinador de Presto utilizando el puerto 8889. Para obtener más información acerca de la interfaz web de Presto, consulte Web Interface en la documentación de Presto.
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versiones de los componentes de la versión 4.8.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	4.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.10.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.3-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.3-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.7.3-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.3-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.3-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.3-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.3-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.3-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.3-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.3-amzn-0	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-7	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-7	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-7	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-7	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-metastore-server	1.0.0-amzn-7	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-7	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.52	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.152.3	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.152.3	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.2	Clientes de línea de comando de Spark.
spark-history-server	1.6.2	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.8.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.8.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.

Clasificaciones	Descripción
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole .properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra .properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.

Clasificaciones	Descripción
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.

Clasificaciones	Descripción
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.8.1

Versiones de las aplicaciones de la versión 4.8.1

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.151	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.3	0.8.3

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 4.8.1

Versiones de los componentes de la versión 4.8.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.9.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.2-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.2-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-7	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-7	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-7	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-7	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-7	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-7	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.51	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.151	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.151	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.2	Clientes de línea de comando de Spark.
spark-history-server	1.6.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.2	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.8.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.8.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.8.0

Versiones de las aplicaciones de la versión 4.8.0

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.148	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 4.8.0

Las siguientes notas de la versión incluyen información sobre la versión 4.8.0 de Amazon EMR. Los cambios son respecto a la versión 4.7.2 de Amazon EMR.

Fecha de lanzamiento: 7 de septiembre de 2016

Actualizaciones

- Se ha actualizado a HBase 1.2.2
- Actualizado a Presto-Sandbox 0.151
- Actualizado a Tez 0.8.4
- Actualizado a Zeppelin-Sandbox 0.6.1

Cambios y mejoras

- Se ha corregido un error en YARN que ApplicationMaster provocaba que intentaran limpiar contenedores que ya no existen porque sus instancias se habían cerrado.
- Se ha corregido la URL de hive-server2 para acciones de Hive2 en los ejemplos de Oozie.
- Se ha añadido soporte para catálogos de Presto adicionales.
- Se han adaptado los parches: HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- Se ha añadido soporte para configuraciones de seguridad, que le permiten crear y aplicar opciones de cifrado con más facilidad. Para obtener más información, consulte Cifrado de datos.

Versiones de los componentes de la versión 4.8.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.9.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.2-amzn-4	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-4	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-4	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-4	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-4	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-4	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-4	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-4	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.2-amzn-4	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-4	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.2	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.2	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.2	HBase cliente de línea de comandos.
hbase-rest-server	1.2.2	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.2	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-7	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-7	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-7	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-7	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-7	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-7	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.51	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.151	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.151	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.2	Clientes de línea de comando de Spark.
spark-history-server	1.6.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.2	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.4	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.6.1	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.8.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.8.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.

Clasificaciones	Descripción
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hiveserver2-site	Cambiar los valores en el archivo hiveserver2-site.xml de Hive Server2.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-me trics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-sit e.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.pro perties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-me trics2-phoenix.properties de Phoenix.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-blackhole	Cambiar los valores en el archivo blackhole.properties de Presto.
presto-connector-cassandra	Cambiar los valores en el archivo cassandra.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
presto-connector-jmx	Cambiar los valores en el archivo jmx.properties de Presto.
presto-connector-kafka	Cambiar los valores en el archivo kafka.properties de Presto.
presto-connector-localfile	Cambiar los valores en el archivo localfile.properties de Presto.
presto-connector-mongodb	Cambiar los valores en el archivo mongodb.properties de Presto.
presto-connector-mysql	Cambiar los valores en el archivo mysql.properties de Presto.
presto-connector-postgresql	Cambiar los valores en el archivo postgresql.properties de Presto.

Clasificaciones	Descripción
presto-connector-raptor	Cambiar los valores en el archivo raptor.properties de Presto.
presto-connector-redis	Cambiar los valores en el archivo redis.properties de Presto.
presto-connector-tpch	Cambiar los valores en el archivo tpch.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Clasificaciones	Descripción
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.7.4

Versiones de las aplicaciones de la versión 4.7.4

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.148	0.147
Spark	1.6.2	1.6.2	1.6.2	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 4.7.4

Esta es una versión de revisión para agregar la autenticación de AWS Signature Version 4 a las solicitudes a Amazon S3. Todas las aplicaciones y componentes son los mismos que en la versión anterior de Amazon EMR.

> **Important**
>
> En esta versión de lanzamiento, Amazon EMR utiliza AWS Signature Version 4 exclusivamente para autenticar solicitudes para Amazon S3. Para obtener más información, consulte Novedades.

Versiones de los componentes de la versión 4.7.4

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.8.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.2-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-3	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.2-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.2.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.1	HBase cliente de línea de comandos.
hbase-rest-server	1.2.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-6	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-6	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-6	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-6	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-6	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server	1.0.0-amzn-6	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.148	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.148	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.2	Clientes de línea de comando de Spark.
spark-history-server	1.6.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.3	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.7.4

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.7.4

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.7.3

Versiones de las aplicaciones de la versión 4.7.3

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,75
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.0	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.147	0.147
Spark	1.6.2	1.6.2	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 4.7.3

Versiones de los componentes de la versión 4.7.3

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.

Componente	Versión	Descripción
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.8.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.2-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.

Componente	Versión	Descripción
hadoop-hdfs-library	2.7.2-amzn-3	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.

Componente	Versión	Descripción
hbase-hmaster	1.2.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.1	HBase cliente de línea de comandos.
hbase-rest-server	1.2.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-6	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-6	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-6	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-6	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-metastore-server	1.0.0-amzn-6	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-6	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica

Componente	Versión	Descripción
presto-coordinator	0.148	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.148	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.2	Clientes de línea de comando de Spark.
spark-history-server	1.6.2	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.3	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.7.3

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.7.3

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop

Clasificaciones	Descripción
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.

Clasificaciones	Descripción
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.

Clasificaciones	Descripción
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.

Clasificaciones	Descripción
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.7.2

Versiones de las aplicaciones de la versión 4.7.2

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
AWS SDK para Java	1.10.75	1.10,75	1.10,75	1.10,27
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.0	0.12.0	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.147	0.147	0.143
Spark	1.6.2	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	-

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 4.7.2

Las siguientes notas de la versión incluyen información sobre la versión 4.7.2 de Amazon EMR.

Fecha de lanzamiento: 15 de julio de 2016

Características

- Actualizado a Mahout 0.12.2
- Actualizado a Presto 0.148
- Actualizado a Spark 1.6.2
- Ahora puede crear un AWSCredentials proveedor para usarlo con EMRFS utilizando un URI como parámetro. Para obtener más información, consulte Crear un AWSCredentials proveedor para EMRFS.
- EMRFS ahora permite a los usuarios configurar un punto de conexión de DynamoDB personalizado para sus metadatos de vista coherente utilizando la propiedad `fs.s3.consistent.dynamodb.endpoint` en `emrfs-site.xml`.
- Se ha añadido un script en `/usr/bin` denominado `spark-example`, que envuelve `/usr/lib/spark/spark/bin/run-example` para que pueda ejecutar los ejemplos directamente. Por ejemplo, para ejecutar el SparkPi ejemplo que viene con la distribución de Spark, puedes ejecutarlo `spark-example SparkPi 100` desde la línea de comandos o usarlo `command-runner.jar` como un paso de la API.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema donde Oozie no tenía `spark-assembly.jar` en la ubicación correcta cuando Spark también está instalado, lo que se traducía en un error al lanzar aplicaciones de Spark con Oozie.
- Se ha solucionado un problema con registros basado en Spark Log4j en contenedores de YARN.

Versiones de los componentes de la versión 4.7.2

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.2.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.1.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.8.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.2-amzn-3	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-3	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-3	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.2-amzn-3	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-3	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-3	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-3	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-3	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-3	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-3	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.

Componente	Versión	Descripción
hbase-region-server	1.2.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.1	HBase cliente de línea de comandos.
hbase-rest-server	1.2.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-6	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-6	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-6	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-6	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-6	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.

Componente	Versión	Descripción
hive-server	1.0.0-amzn-6	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.2	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.148	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.148	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.2	Clientes de línea de comando de Spark.
spark-history-server	1.6.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.3	La aplicación YARN de tez y bibliotecas.
webserver	2.4.23	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.

Componente	Versión	Descripción
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.7.2

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.7.2

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hadoop-ssl-server	Cambiar la configuración del servidor ssl de Hadoop
hadoop-ssl-client	Cambiar la configuración del cliente ssl de Hadoop
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.

Clasificaciones	Descripción
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.

Clasificaciones	Descripción
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.

Clasificaciones	Descripción
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.

Clasificaciones	Descripción
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.7.1

Versiones de las aplicaciones de la versión 4.7.1

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
AWS SDK para Java	1.10.75	1.10,75	1.10,27	1.10,27
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.0	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.12.0	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.147	0.143	0.143
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

Notas de la versión 4.7.1

Las siguientes notas de la versión incluyen información sobre la versión 4.7.1 de Amazon EMR.

Fecha de lanzamiento: 10 de junio de 2016

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un error que ampliaba el tiempo de startup de clústeres lanzadas en una VPC con subredes privadas. El error únicamente afectaba a clústeres lanzados con la versión 4.7.0 de Amazon EMR.
- Se ha corregido un error que gestionaba incorrectamente el listado de archivos en Amazon EMR para los clústeres iniciados con la versión 4.7.0 de Amazon EMR.

Versiones de los componentes de la versión 4.7.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.7.1	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.2-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.2-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.1	HBase cliente de línea de comandos.
hbase-rest-server	1.2.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-5	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-5	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-5	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-5	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-5	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-5	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.0	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.147	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.147	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.1	Clientes de línea de comando de Spark.
spark-history-server	1.6.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.3	La aplicación YARN de tez y bibliotecas.
webserver	2.4.18	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.7.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.7.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.

Clasificaciones	Descripción
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.

Clasificaciones	Descripción
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Clasificaciones	Descripción
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.7.0

Versiones de las aplicaciones de la versión 4.7.0

Esta versión incluye las siguientes aplicaciones: Ganglia,, HBase, HCatalog, Hadoop, Hive, Hue,, Mahout, Oozie-Sandbox, Phoenix, Pig, Presto-Sandbox,, SparkSqoop-SandboxTez, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
AWS SDK para Java	1.10.75	1.10,27	1.10,27	1.10,27
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.0	1.2.0	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.143	0.143	0.140
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	-

Notas de la versión 4.7.0

> **Important**
>
> Amazon EMR 4.7.0 está obsoleto. Utilice Amazon EMR 4.7.1 o una versión posterior en su lugar.

Fecha de lanzamiento: 2 de junio de 2016

Características

- Se ha añadido Apache Phoenix 4.7.0
- Se ha añadido Apache Tez 0.8.3
- Se ha actualizado a HBase 1.2.1
- Actualizado a Mahout 0.12.0
- Actualizado a Presto 0.147
- Se actualizó AWS SDK para Java a la versión 1.10.75
- La marca final se ha eliminado de la propiedad `mapreduce.cluster.local.dir` en `mapred-site.xml` para permitir a los usuarios ejecutar Pig en modo local.
- Los controladores JDBC de Amazon Redshift están disponibles en el clúster

  Los controladores JDBC de Amazon Redshift ahora se incluyen en `/usr/share/aws/redshift/jdbc`. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` es el controlador de Amazon Redshift compatible con JDBC 4.1 y `/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar` es el controlador de Amazon Redshift compatible con JDBC 4.0. Para más información, consulte Configurar una conexión JDBC en la Guía de administración de Amazon Redshift.
- Java 8

  Excepto para Presto, OpenJDK 1.7 es el JDK predeterminado para todas las aplicaciones. No obstante, tanto OpenJDK 1.7 como 1.8 están instalados. Para obtener más información sobre cómo definir `JAVA_HOME` para las aplicaciones, consulte Configuración de aplicaciones para utilizar Java 8.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema de kernel que afectaba notablemente a los volúmenes de HDD de rendimiento optimizado (st1) de EBS para Amazon EMR en emr-4.6.0.
- Se ha corregido un problema donde un clúster fallaría si alguna de la zona de cifrado de HDFS se especifica sin elegir Hadoop como aplicación.
- Se ha cambiado la política de escritura de HDFS predeterminada `RoundRobin` a `AvailableSpaceVolumeChoosingPolicy`. Algunos volúmenes no se utilizaron correctamente con la RoundRobin configuración, lo que provocó fallos en los nodos principales y un HDFS poco fiable.

- Se ha corregido un problema con la CLI de EMRFS que provocaría una excepción al crear la tabla de metadatos de DynamoDB predeterminada para vistas coherentes.
- Se ha corregido un problema de bloqueo en EMRFS que ocurría potencialmente durante las operaciones de copia y cambio de nombre de multipartes.
- Se ha corregido un problema con EMRFS que provocaba que el CopyPart tamaño predeterminado fuera de 5 MB. El valor predeterminado se ha fijado correctamente ahora en 128 MB.
- Se ha corregido un problema con la configuración de upstart de Zeppelin que potencialmente impedía parar el servicio.
- Se ha corregido un problema con Spark y Zeppelin, que impedía que utilizara el esquema de URI de `s3a://` porque `/usr/lib/hadoop/hadoop-aws.jar` no se cargó correctamente en la classpath respectiva.
- Se ha adaptado HUE-2484.
- Se respaldó una confirmación de Hue 3.9.0 (no existe JIRA) para solucionar un problema con la muestra del navegador. HBase
- Se ha adaptado HIVE-9073.

Versiones de los componentes de la versión 4.7.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.1.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.2.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.4.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.7.1	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.2-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resour cemanager	2.7.2-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hadoop-yarn-timeline-server	2.7.2-amzn-2	Servicio para recuperar información actual e histórica para aplicaciones de YARN.
hbase-hmaster	1.2.1	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.1	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.1	HBase cliente de línea de comandos.
hbase-rest-server	1.2.1	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.1	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-5	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-5	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-5	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-5	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-5	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-5	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-7	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.12.0	Biblioteca para machine learning.
mysql-server	5.5.46	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
phoenix-library	4.7.0- 1.2 HBase	Las bibliotecas de Phoenix para servidor y cliente
phoenix-query-server	4.7.0- 1.2 HBase	Un servidor ligero que proporciona acceso a JDBC, así como acceso de formato Protocol Buffers y JSON al API de Avatica
presto-coordinator	0.147	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.147	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.1	Clientes de línea de comando de Spark.
spark-history-server	1.6.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
tez-on-yarn	0.8.3	La aplicación YARN de tez y bibliotecas.
webserver	2.4.18	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.7.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.7.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.

Clasificaciones	Descripción
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.

Clasificaciones	Descripción
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
phoenix-hbase-metrics	Cambiar los valores en el archivo hadoop-metrics2-hbase.properties de Phoenix.
phoenix-hbase-site	Cambiar los valores en el archivo hbase-site.xml de Phoenix.
phoenix-log4j	Cambiar los valores en el archivo log4j.properties de Phoenix.
phoenix-metrics	Cambiar los valores en el archivo hadoop-metrics2-phoenix.properties de Phoenix.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.

Clasificaciones	Descripción
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
tez-site	Cambiar los valores en el archivo tez-site.xml de Tez.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Clasificaciones	Descripción
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.6.1

Versiones de las aplicaciones de la versión 4.6.1

Esta versión incluye las siguientes aplicaciones: Ganglia,, HBase, HCatalog, Hadoop, Hive,, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, SparkSqoop-Sandbox, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
AWS SDK para Java	1.10.27	1.10,27	1.10,27	1.10,27
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	1.2.0	-	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.	2.7.	2.7.	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.143	0.140	0.136
Spark	1.6.1	1.6.1	1.6.1	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	-	-

Notas de la versión 4.6.1

Versiones de los componentes de la versión 4.6.1

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que

pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.1.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.3.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.6.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregació

Componente	Versión	Descripción
		n de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.2-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider

Componente	Versión	Descripción
hadoop-mapred	2.7.2-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.2-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hbase-hmaster	1.2.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.0	HBase cliente de línea de comandos.
hbase-rest-server	1.2.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.

Componente	Versión	Descripción
hcatalog-server	1.0.0-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-4	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-6	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.11.1	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.

Componente	Versión	Descripción
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
presto-coordinator	0.143	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.143	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.1	Clientes de línea de comando de Spark.
spark-history-server	1.6.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
webserver	2.4	Servidor HTTP de Apache.

Componente	Versión	Descripción
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.6.1

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.6.1

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.

Clasificaciones	Descripción
hbase-env	Cambie los valores en HBase el entorno.
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.

Clasificaciones	Descripción
hive-log4j	Cambiar los valores en el archivo hive-log4 .properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
yarn-env	Cambiar los valores en el entorno de YARN.

Clasificaciones	Descripción
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.6.0

Versiones de las aplicaciones de la versión 4.6.0

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HBase, HCatalog, Hadoop, Hive,, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox, SparkSqoop-Sandbox, Zeppelin-Sandboxy. ZooKeeper-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
AWS SDK para Java	1.10.27	1.10,27	1.10,27	1.10,27
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	-	-	-
HCatalog	1.0.0	1.0.0	1.0.0	-
Hadoop	2.7.	2.7.	2.7.1	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
Mahout	0.11.1	0.11.1	0.11.1	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.140	0.136	0.130
Spark	1.6.1	1.6.1	1.6.0	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	-	-	-

Notas de la versión 4.6.0

Las siguientes notas de la versión incluyen información sobre la versión 4.6.0 de Amazon EMR.

- Se ha agregado HBase 1.2.0
- Se ha añadido Zookeeper-Sandbox 3.4.8
- Actualizado a Presto-Sandbox 0.143
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.03.0. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/.
- Problema que afecta a los tipos de volúmenes de EBS HDD con velocidad optimizada (st1)

  Un problema en el kernel de Linux versiones 4.2 y superiores afecta notablemente al rendimiento en volúmenes de EBS HDD con velocidad optimizada (st1) para EMR. Esta versión (emr-4.6.0) utiliza el kernel versión 4.4.5 y, por tanto, se ve afectada. Por lo tanto, le recomendamos no utilizar emr-4.6.0 si desea utilizar los volúmenes de EBS st1. Puede utilizar emr-4.5.0 o versiones anteriores de Amazon EMR sin que afecten a st1. Además, proporcionamos la corrección en futuras versiones.
- Valores predeterminados de Python

  Python 3.4 se instala ahora de forma predeterminada, pero Python 2.7 sigue siendo el valor predeterminado del sistema. Puede configurar Python 3.4 como predeterminado del sistema mediante una acción de arranque; puede usar la API de configuración para establecer la exportación de PYSPARK_PYTHON `/usr/bin/python3.4` en la `spark-env` clasificación para que afecte a la versión de Python utilizada por. PySpark
- Java 8

  Excepto para Presto, OpenJDK 1.7 es el JDK predeterminado para todas las aplicaciones. No obstante, tanto OpenJDK 1.7 como 1.8 están instalados. Para obtener más información sobre cómo definir `JAVA_HOME` para las aplicaciones, consulte Configuración de aplicaciones para utilizar Java 8.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema donde el aprovisionamiento de aplicaciones en ocasiones fallaría aleatoriamente debido a una contraseña generada.
- Anteriormente, `mysqld` estaba instalado en todos los nodos. Ahora, solo está instalado en la instancia principal y únicamente si la aplicación elegida incluye `mysql-server` como

componente. Actualmente, las siguientes aplicaciones incluyen el `mysql-server` componente: Hive, Hue HCatalog, Presto-Sandbox y Sqoop-Sandbox.

- Se ha cambiado `yarn.scheduler.maximum-allocation-vcores` a 80 desde el valor predeterminado de 32, que corrige un problema introducido en emr-4.4.0 que se produce sobre todo con Spark al utilizar la opción `maximizeResourceAllocation` en un clúster cuyo tipo de instancia secundaria es uno de los pocos tipos de instancia grandes que tienen los vcores de YARN definidos en un valor superior a 32; es decir, c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge o m4.10xlarge se han visto afectados por este problema.
- s3-dist-cp ahora utiliza EMRFS para todas las nominaciones de Amazon S3 y ya no hay fases en un directorio temporal de HDFS.
- Se ha corregido un problema con la gestión de excepciones para las cargas multiparte de cifrado del lado del cliente.
- Se ha agregado una opción para permitir a los usuarios cambiar la clase de almacenamiento de Amazon S3. De forma predeterminada, este valor es `STANDARD`. El valor de clasificación de configuración `emrfs-site` es `fs.s3.storageClass` y los valores posibles son `STANDARD`, `STANDARD_IA` y `REDUCED_REDUNDANCY`. Para más información sobre las clases de almacenamiento, consulte Clases de almacenamiento en la Guía del usuario de Amazon Simple Storage Service.

Versiones de los componentes de la versión 4.6.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.1.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.3.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.6.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.2-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.2-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hbase-hmaster	1.2.0	Servicio para un HBase clúster responsable de la coordinación de las regiones y la ejecución de los comandos administrativos.
hbase-region-server	1.2.0	Servicio para prestar servicio a una o más HBase regiones.
hbase-client	1.2.0	HBase cliente de línea de comandos.
hbase-rest-server	1.2.0	Servicio que proporciona un punto final RESTful HTTP para. HBase
hbase-thrift-server	1.2.0	Servicio que proporciona un punto final de Thrift para HBase.
hcatalog-client	1.0.0-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	1.0.0-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-4	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-6	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.11.1	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.

Componente	Versión	Descripción
presto-coordinator	0.143	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.143	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.1	Clientes de línea de comando de Spark.
spark-history-server	1.6.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
webserver	2.4	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.

Componente	Versión	Descripción
zookeeper-server	3.4.8	Servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo.
zookeeper-client	3.4.8	ZooKeeper cliente de línea de comandos.

Clasificaciones de configuración de la versión 4.6.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.6.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.pro perties de Hadoop.
hbase-env	Cambie los valores en HBase el entorno.

Clasificaciones	Descripción
hbase-log4j	Cambie los valores en el archivo HBase hbase-log4j.properties.
hbase-metrics	Cambie los valores en el archivo hadoop-metrics2-hbaase.properties. HBase
hbase-policy	Cambie los valores en HBase el archivo hbase-policy.xml.
hbase-site	Cambia los valores en HBase el archivo hbase-site.xml.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.

Clasificaciones	Descripción
hive-log4j	Cambiar los valores en el archivo hive-log4 .properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-si te.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.

Clasificaciones	Descripción
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
yarn-env	Cambiar los valores en el entorno de YARN.

Clasificaciones	Descripción
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.
zookeeper-config	Cambie los valores en el ZooKeeper archivo zoo.cfg.
zookeeper-log4j	Cambia los valores en ZooKeeper el archivo de propiedades.

Amazon EMR, versión 4.5.0

Versiones de las aplicaciones de la versión 4.5.0

Esta versión incluye las siguientes aplicaciones: Ganglia,,, HCatalog, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox,, PigPresto-SandboxSpark, Sqoop-Sandboxy. Zeppelin-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
AWS SDK para Java	1.10.27	1.10,27	1.10,27	1.10,27

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.6.0
HBase	-	-	-	-
HCatalog	1.0.0	1.0.0	-	-
Hadoop	2.7.	2.7.1	2.7.1	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.0	0.11.0
Oozie	-	-	-	-

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.140	0.136	0.130	0.125
Spark	1.6.1	1.6.0	1.6.0	1.5.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.5	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas de la versión 4.5.0

Las siguientes notas de la versión incluyen información sobre la versión 4.5.0 de Amazon EMR.

Fecha de lanzamiento: 4 de abril de 2016

Características

- Actualizado a Spark 1.6.1
- Actualizado a Hadoop 2.7.2
- Actualizado a Presto 0.140
- Se ha agregado AWS KMS compatibilidad con para el cifrado del lado del servidor de Amazon S3.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema donde los servidores de MySQL y Apache no se iniciaban después de reiniciar un nodo.
- Se ha corregido un problema donde IMPORT no funcionaba correctamente con tablas no particionadas almacenadas en Amazon S3
- Se ha corregido un problema en Presto donde requiere que el directorio intermedio sea `/mnt/tmp` en lugar de `/tmp` al escribir en tablas de Hive.

Versiones de los componentes de la versión 4.5.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.1.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.2.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.5.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.2-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.2-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.2-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.2-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.2-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.2-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.2-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.2-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.2-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hcatalog-client	1.0.0-amzn-4	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-4	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.
hcatalog-webhcat-server	1.0.0-amzn-4	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-4	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-4	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-4	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-5	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop

Componente	Versión	Descripción
mahout-client	0.11.1	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.
presto-coordinator	0.140	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.140	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.1	Clientes de línea de comando de Spark.
spark-history-server	1.6.1	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.1	Motor de ejecución en memoria para YARN.

Componente	Versión	Descripción
spark-yarn-slave	1.6.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
webserver	2.4	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.

Clasificaciones de configuración de la versión 4.5.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.5.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.

Clasificaciones	Descripción
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.

Clasificaciones	Descripción
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Amazon EMR, versión 4.4.0

Versiones de las aplicaciones de la versión 4.4.0

Esta versión incluye las siguientes aplicaciones: Ganglia, HCatalog,, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-SandboxSparkSqoop-Sandbox, y. Zeppelin-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
AWS SDK para Java	1.10.27	1.10,27	1.10,27	No registrado
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.6.0	-
HBase	-	-	-	-
HCatalog	1.0.0	-	-	-
Hadoop	2.7.1	2.7.1	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.0	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.0.1
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.136	0.130	0.125	0.119
Spark	1.6.0	1.6.0	1.5.2	1.5.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.5	0.5.5	0.6.0-SNAPSHOT
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas de la versión 4.4.0

Las siguientes notas de la versión incluyen información sobre la versión 4.4.0 de Amazon EMR.

Fecha de lanzamiento: 14 de marzo de 2016

Características

- Se ha agregado HCatalog 1.0.0
- Se ha añadido Sqoop-Sandbox 1.4.6
- Actualizado a Presto 0.136
- Actualizado a Zeppelin 0.5.6
- Actualizado a Mahout 0.11.1
- `dynamicResourceAllocation` habilitado de forma predeterminada.
- Se ha añadido una tabla de todas las clasificaciones de configuración para esta versión. Para obtener más información, consulte la tabla de clasificaciones de configuración de Configuración de aplicaciones.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un error donde la `maximizeResourceAllocation` configuración no reservaría memoria suficiente para ApplicationMaster daemons YARN.

- Se ha corregido un problema encontrado con una DNS personalizada. Si alguna entrada en `resolve.conf` precede a las entradas personalizadas proporcionada, entonces las entradas personalizadas no se pueden resolver. Este comportamiento se veía afectado por clústeres en una VPC donde el servidor de nombres de VPC predeterminado se inserta como entrada superior en `resolve.conf`.
- Se ha corregido un problema donde la versión predeterminada de Python se traslada a la versión 2.7 y boto no se ha instalado para dicha versión.
- Se ha corregido un error donde los contenedores de YARN y las aplicaciones de Spark generarían un archivo único de base de datos round robin (rrd) de Ganglia, que daría lugar a que se llenara el primer disco asociado a la instancia. Debido a esta solución, las métricas en el nivel de contenedor de YARN se han deshabilitado y las métricas en el nivel de aplicación de Spark se han deshabilitado.
- Se ha corregido un problema en el insertador de registros que eliminaría todas las carpetas con registro vacío. El resultado era que la CLI de Hive no podía realizar el registro porque el insertador de registros eliminaba la carpeta `user` vacía en `/var/log/hive`.
- Se ha corregido un problema que afectaba a las importaciones de Hive, que afectaba a las particiones y daba lugar a un error durante la importación.
- Se ha corregido un problema donde EMRFS y s3-dist-cp no gestionaban de forma adecuada los nombres de bucket que contenían puntos.
- Se ha cambiado un comportamiento en EMRFS, de forma que los buckets con control de versiones habilitado en el archivo de marcador `_$folder$` no se creaba de forma continua, lo que podría contribuir a mejorar el rendimiento de los buckets con control de versiones habilitado.
- Se ha cambiado el comportamiento en EMRFS de forma que no utiliza archivos de instrucciones, excepto en aquellos casos en los que el cifrado del lado del cliente está habilitado. Si desea eliminar archivos de instrucciones al utilizar el cifrado del lado cliente, puede establecer la propiedad de emrfs-site.xml, `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled`, en true.
- Se ha cambiado la agregación de registros de YARN para conservar los registros en el destino agregación durante dos días. El destino predeterminado es el almacenamiento de HDFS de su clúster. Si desea cambiar esta duración, cambie el valor de `yarn.log-aggregation.retain-seconds` utilizando la clasificación de configuración `yarn-site` al crear el clúster. Como siempre, puede guardar los registros de las aplicaciones en Amazon S3 utilizando el parámetro `log-uri` al crear el clúster.

Parches aplicados

- HIVE-9655
- HIVE-9183
- HADOOP-12810

Versiones de los componentes de la versión 4.4.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.1.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-s3-dist-cp	2.2.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.4.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.7.1-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.1-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.1-amzn-1	Biblioteca y cliente de línea de comandos HDFS

Componente	Versión	Descripción
hadoop-hdfs-namenode	2.7.1-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.1-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.1-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.1-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.1-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.7.1-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hcatalog-client	1.0.0-amzn-3	El cliente de línea de comando "hcat" para manipular hcatalog-server.
hcatalog-server	1.0.0-amzn-3	Prestación de servicios HCatalog, una capa de administración de tablas y almacenamiento para aplicaciones distribuidas.

Componente	Versión	Descripción
hcatalog-webhcat-server	1.0.0-amzn-3	Punto final HTTP que proporciona una interfaz REST para HCatalog.
hive-client	1.0.0-amzn-3	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-3	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-3	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-5	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.11.1	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.

Componente	Versión	Descripción
presto-coordinator	0.136	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.136	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.0	Clientes de línea de comando de Spark.
spark-history-server	1.6.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
sqoop-client	1.4.6	Cliente de línea de comando de Apache Sqoop.
webserver	2.4	Servidor HTTP de Apache.
zeppelin-server	0.5.6-incubating	Bloc de notas basado en web que permite el análisis de datos interactivo.

Clasificaciones de configuración de la versión 4.4.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.4.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hcatalog-env	Cambie los valores en HCatalog el entorno.
hcatalog-server-jndi	Cambie los valores en HCatalog jndi.properties.
hcatalog-server-proto-hive-site	Cambie los valores en .xml HCatalog. proto-hive-site
hcatalog-webhcat-env	Cambie los valores en el entorno HCat de la HCatalog Web.
hcatalog-webhcat-log4j	Cambie los valores en el HCatalog HCat log4j.properties de la Web.

Clasificaciones	Descripción
hcatalog-webhcat-site	Cambia los valores en el archivo webhcat-site.xml HCat de la HCatalog Web.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.

Clasificaciones	Descripción
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-def aults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.pro perties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.p roperties de Spark.
sqoop-env	Cambiar los valores en el entorno de Sqoop.

Clasificaciones	Descripción
sqoop-oraoop-site	Cambia los valores en el archivo oraoop-site.xml de OraOop Sqoop.
sqoop-site	Cambiar los valores en el archivo sqoop-site.xml de Sqoop.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Amazon EMR, versión 4.3.0

Versiones de las aplicaciones de la versión 4.3.0

Esta versión incluye las siguientes aplicaciones: Ganglia, Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, PigPresto-SandboxSpark, y. Zeppelin-Sandbox

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK para Java	1.10.27	1.10,27	No registrado	No registrado
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas de la versión 4.3.0

Las siguientes notas de la versión incluyen información sobre la versión 4.3.0 de Amazon EMR.

Fecha de lanzamiento: 19 de enero de 2016

Características

- Actualizado a Hadoop 2.7.1
- Actualizado a Spark 1.6.0
- Actualizado Ganglia a 3.7.2
- Actualizado Presto a 0.130
- Amazon EMR ha hecho algunos cambios en `spark.dynamicAllocation.enabled` cuando se define en true; su valor predeterminado es false. Cuando se define en true, esto afecta a los valores predeterminados del ajuste `maximizeResourceAllocation`:
  - Si `spark.dynamicAllocation.enabled` se ha definido en true, `spark.executor.instances` no ha sido definido por `maximizeResourceAllocation`.
  - El ajuste `spark.driver.memory` se configura ahora en función de los tipos de instancia en el clúster de forma similar a cómo se define `spark.executors.memory`. Sin embargo, dado que la aplicación del controlador Spark puede ejecutarse en la instancia principal o en una de las instancias secundarias (por ejemplo, en un cliente de YARN y modos de clústeres, respectivamente), el ajuste `spark.driver.memory` se basa en el tipo de instancia del tipo de instancia más pequeño de estos dos grupos de instancias.
  - El ajuste `spark.default.parallelism` se define ahora como el doble del número de núcleos de CPU disponibles para contenedores de YARN. En las versiones anteriores, era la mitad de dicho valor.
  - Los cálculos de la sobrecarga de memoria reservada para procesos de Spark YARN se han ajustado para ser más precisos, lo que se traduce en un pequeño aumento en la cantidad total de memoria disponible para Spark (es decir, `spark.executor.memory`).

Problemas conocidos resueltos de las versiones anteriores

- La agregación de registros de YARN ahora está habilitada de forma predeterminada.
- Se ha corregido un problema por el que los registros no se enviaban a un bucket de registros de Amazon S3 del clúster cuando la agregación de registros de YARN estaba habilitada.

- Los tamaños de contenedor de YARN ahora tienen un nuevo mínimo de 32 en todos los tipos de nodos.
- Se ha corregido un problema con Ganglia que provocaba E/S de disco excesiva en el nodo principal en clústeres grandes.
- Se ha corregido un problema que impedía que los registros de las aplicaciones se enviaran a Amazon S3 cuando se cierra un clúster.
- Se ha corregido un problema en la CLI de EMRFS que provocaba que determinados comandos devolvieran error.
- Se ha corregido un problema con Zeppelin que impedía que las dependencias se cargaran en la base. SparkContext
- Se ha corregido un problema que se producía al intentar un ajuste de tamaño para añadir instancias.
- Se ha corregido un problema en Hive donde CREATE TABLE AS SELECT realiza llamadas de lista excesivas a Amazon S3.
- Se ha corregido un problema donde clústeres grandes no aprovisionarían correctamente cuando estaban instalados Hue, Oozie y Ganglia.
- Se ha corregido un problema en s3-dist-cp que devolvería un código de salida cero incluso si ha devuelto un error.

Versiones de los componentes de la versión 4.3.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.1.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.1.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.3.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.7.2	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.7.2	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.7.1	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.

Componente	Versión	Descripción
hadoop-client	2.7.1-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.7.1-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.7.1-amzn-0	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.7.1-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.7.1-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.7.1-amzn-0	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.7.1-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.7.1-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.7.1-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hive-client	1.0.0-amzn-2	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-2	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-2	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-5	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.11.0	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.

Componente	Versión	Descripción
presto-coordinator	0.130	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.130	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.6.0	Clientes de línea de comando de Spark.
spark-history-server	1.6.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.6.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.6.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
webserver	2.4	Servidor HTTP de Apache.
zeppelin-server	0.5.5-incubating-amzn-1	Bloc de notas basado en web que permite el análisis de datos interactivo.

Clasificaciones de configuración de la versión 4.3.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.3.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue

Clasificaciones	Descripción
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j.properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.

Clasificaciones	Descripción
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.properties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.properties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Amazon EMR, versión 4.2.0

Versiones de las aplicaciones de la versión 4.2.0

Esta versión incluye las siguientes aplicaciones: GangliaHadoop, Hive, Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark, y Zeppelin-Sandbox.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK para Java	1.10.27	1.10,27	No registrado	No registrado
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas de la versión 4.2.0

Las siguientes notas de la versión incluyen información sobre la versión 4.2.0 de Amazon EMR.

Fecha de lanzamiento: 18 de noviembre de 2015

Características

- Se ha añadido compatibilidad con Ganglia
- Actualizado a Spark 1.5.2
- Actualizado a Presto 0.125
- Actualizado Oozie a 4.2.0
- Actualizado Zeppelin a 0.5.5
- Actualizado a la versión AWS SDK para Java 1.10.27

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema con la CLI de EMRFS, donde no se utilizaba el nombre de la tabla de metadatos predeterminado.
- Se ha corregido un problema que se presentaba al utilizar tablas respaldadas por ORC en Amazon S3.

- Se ha corregido un problema encontrado con un error de coincidencia de versión de Python en la configuración de Spark.
- Se ha corregido un error cuando falla la notificación del estado de un nodo de YARN debido a problemas de DNS para clústeres en una VPC.
- Se ha corregido un error que se presentaba cuando YARN retiraba nodos, lo que daba lugar a aplicaciones que no respondían o a la incapacidad de programar nuevas aplicaciones.
- Se ha corregido un error que se encuentra cuando los clústeres terminan con el estado TIMED_OUT_STARTING.
- Se ha corregido un error que se encuentra al incluir la dependencia de EMRFS Scala en otras versiones. Se ha eliminado la dependencia de Scala.

Versiones de los componentes de la versión 4.2.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.1.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.0.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.2.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.
ganglia-monitor	3.6.0	Agente de Ganglia incrustado para aplicaciones del ecosistema de Hadoop junto con el agente de monitorización de Ganglia.
ganglia-metadata-collector	3.6.0	Recopilador de metadatos de Ganglia para agregación de métricas a partir de los agentes de monitorización de Ganglia.
ganglia-web	3.5.10	Aplicación web para visualizar las métricas recopiladas por el recopilador de metadatos de Ganglia.
hadoop-client	2.6.0-amzn-2	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".

Componente	Versión	Descripción
hadoop-hdfs-datanode	2.6.0-amzn-2	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.6.0-amzn-2	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.6.0-amzn-2	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.6.0-amzn-2	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.6.0-amzn-2	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.6.0-amzn-2	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.6.0-amzn-2	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.6.0-amzn-2	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hive-client	1.0.0-amzn-1	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-metastore-server	1.0.0-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-5	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.11.0	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
oozie-client	4.2.0	Cliente de línea de comando de Oozie.
oozie-server	4.2.0	Servicio para aceptar solicitud es de flujo de trabajo de Oozie.
presto-coordinator	0.125	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.125	Service para ejecutar partes de una consulta.

Componente	Versión	Descripción
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.5.2	Clientes de línea de comando de Spark.
spark-history-server	1.5.2	IU web para la visualización de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.5.2	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.5.2	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
webserver	2.4	Servidor HTTP de Apache.
zeppelin-server	0.5.5-incubating-amzn-0	Bloc de notas basado en web que permite el análisis de datos interactivo.

Clasificaciones de configuración de la versión 4.2.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.2.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.

Clasificaciones	Descripción
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.

Clasificaciones	Descripción
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.
presto-connector-hive	Cambiar los valores en el archivo hive.prop erties de Presto.
spark	Configuración de Amazon EMR mantenida para Apache Spark.

Clasificaciones	Descripción
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
spark-metrics	Cambiar los valores en el archivo metrics.properties de Spark.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Amazon EMR, versión 4.1.0

Versiones de las aplicaciones de la versión 4.1.0

Esta versión incluye las siguientes aplicaciones: HadoopHive, Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark, y Zeppelin-Sandbox.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK para Java	1.10.27	1.10,27	No registrado	No registrado
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas de la versión 4.1.0

Versiones de los componentes de la versión 4.1.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.1.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.0.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.1.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
hadoop-client	2.6.0-amzn-1	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.6.0-amzn-1	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-library	2.6.0-amzn-1	Biblioteca y cliente de línea de comandos HDFS
hadoop-hdfs-namenode	2.6.0-amzn-1	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.6.0-amzn-1	Punto de enlace HTTP para operaciones HDFS.
hadoop-kms-server	2.6.0-amzn-1	Servidor de administración de claves criptográficas basado en la API de Hadoop. KeyProvider
hadoop-mapred	2.6.0-amzn-1	MapReduce bibliotecas de motores de ejecución para ejecutar una aplicación. MapReduce
hadoop-yarn-nodemanager	2.6.0-amzn-1	Servicio de YARN para la administración de contenedo res en un nodo individual.

Componente	Versión	Descripción
hadoop-yarn-resourcemanager	2.6.0-amzn-1	Servicio de YARN para la asignación y administración de recursos de clúster y aplicaciones distribuidas.
hive-client	1.0.0-amzn-1	Cliente de línea de comando de Hive.
hive-metastore-server	1.0.0-amzn-1	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-1	Servicio para aceptar consultas de Hive como solicitudes web.
hue-server	3.7.1-amzn-4	Aplicación web para analizar datos con aplicaciones del ecosistema de Hadoop
mahout-client	0.11.0	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
oozie-client	4.0.1	Cliente de línea de comando de Oozie.
oozie-server	4.0.1	Servicio para aceptar solicitudes de flujo de trabajo de Oozie.

Componente	Versión	Descripción
presto-coordinator	0.119	Servicio para aceptar las consultas y administrar la ejecución de consultas entre presto-workers.
presto-worker	0.119	Service para ejecutar partes de una consulta.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.5.0	Clientes de línea de comando de Spark.
spark-history-server	1.5.0	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.5.0	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.5.0	Bibliotecas de Apache Spark necesarias para esclavos de YARN.
zeppelin-server	0.6.0-incubating-SNAPSHOT	Bloc de notas basado en web que permite el análisis de datos interactivo.

Clasificaciones de configuración de la versión 4.1.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.1.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hdfs-encryption-zones	Configurar zonas de cifrado de HDFS.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
hue-ini	Cambiar los valores en el archivo ini de Hue
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
hadoop-kms-acls	Cambiar los valores en el archivo kms-acls.xml de Hadoop.
hadoop-kms-env	Cambiar los valores en el entorno de Hadoop KMS.
hadoop-kms-log4j	Cambiar los valores en el archivo kms-log4j .properties de Hadoop.
hadoop-kms-site	Cambiar los valores en el archivo kms-site.xml de Hadoop.
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-si te.xml de la MapReduce aplicación.
oozie-env	Cambiar los valores en el entorno de Oozie.
oozie-log4j	Cambiar los valores en el archivo oozie-log 4j.properties de Oozie.
oozie-site	Cambiar los valores en el archivo oozie-site.xml de Oozie.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.pro perties de Pig.
presto-log	Cambiar los valores en el archivo log.properties de Presto.
presto-config	Cambiar los valores en el archivo config.pr operties de Presto.

Clasificaciones	Descripción
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.
zeppelin-env	Cambiar los valores en el entorno de Zeppelin.

Amazon EMR, versión 4.0.0

Versiones de las aplicaciones de la versión 4.0.0

Esta versión incluye las siguientes aplicaciones: HadoopHiveMahout, Pig, y Spark.

En la siguiente tabla se enumeran las versiones de la aplicación disponibles en esta versión de Amazon EMR y las versiones de la aplicación en las tres versiones anteriores de Amazon EMR (cuando corresponda).

Para ver un historial completo de las versiones de la aplicación disponibles para cada versión de Amazon EMR, consulte los temas siguientes:

- Versiones de las aplicaciones en las versiones 7.x de Amazon EMR
- Versiones de la aplicación en las versiones 6.x de Amazon EMR
- Versiones de la aplicación en las versiones 5.x de Amazon EMR
- Versiones de la aplicación en las versiones 4.x de Amazon EMR

Información sobre la versión de la aplicación

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS SDK para Java	1.10.27	1.10,27	No registrado	No registrado
Python	No registrado	No registrado	No registrado	No registrado
Scala	No registrado	No registrado	No registrado	No registrado
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

Notas de la versión 4.0.0

Versiones de los componentes de la versión 4.0.0

A continuación, se muestran los componentes que Amazon EMR instala con esta versión. Algunos se instalan como parte de paquetes de aplicación de Big Data. Otros son exclusivos de Amazon EMR y se instalan para ciertos procesos y características del sistema. Normalmente, estos componentes comienzan con `emr` o `aws`. Normalmente, los paquetes de aplicación de macrodatos de la versión más reciente de Amazon EMR son las versiones más recientes que pueden encontrarse en la comunidad. Intentamos que las versiones de la comunidad estén disponibles en Amazon EMR lo más rápido posible.

Algunos componentes de Amazon EMR son distintos de las versiones que se encuentran en la comunidad. Estos componentes tienen una etiqueta de versión con el formato *`CommunityVersion`*`-amzn-`*`EmrVersion`*. La *`EmrVersion`* empieza por 0. Por ejemplo, si un componente de la comunidad de código abierto llamado `myapp-component` con la versión 2.2 se ha modificado tres veces para incluirlo en diferentes versiones de lanzamiento de Amazon EMR, la versión que se mostrará será `2.2-amzn-2`.

Componente	Versión	Descripción
emr-ddb	3.0.0	Conector de Amazon DynamoDB para aplicaciones del ecosistema de Hadoop.
emr-goodies	2.0.0	Bibliotecas especialmente prácticas para el ecosistema de Hadoop.
emr-kinesis	3.0.0	Conector de Amazon Kinesis para aplicaciones del ecosistema de Hadoop.
emr-s3-dist-cp	2.0.0	Aplicación de copia distribuida optimizada para Amazon S3.
emrfs	2.0.0	Conector de Amazon S3 para aplicaciones del ecosistema de Hadoop.

Componente	Versión	Descripción
hadoop-client	2.6.0-amzn-0	Los clientes de línea de comando de Hadoop como, por ejemplo "hdfs", "hadoop" o "yarn".
hadoop-hdfs-datanode	2.6.0-amzn-0	Servicio de nivel de nodos de HDFS para el almacenamiento de bloques.
hadoop-hdfs-namenode	2.6.0-amzn-0	Servicio de HDFS para realizar un seguimiento de nombres de archivo y bloquear ubicaciones.
hadoop-httpfs-server	2.6.0-amzn-0	Punto de enlace HTTP para operaciones HDFS.
hadoop-mapred	2.6.0-amzn-0	MapReduce bibliotecas de motores de ejecución para ejecutar una MapReduce aplicación.
hadoop-yarn-nodemanager	2.6.0-amzn-0	Servicio de YARN para la administración de contenedo res en un nodo individual.
hadoop-yarn-resour cemanager	2.6.0-amzn-0	Servicio de YARN para la asignación y administración de recursos de clúster y aplicacio nes distribuidas.
hive-client	1.0.0-amzn-0	Cliente de línea de comando de Hive.

Componente	Versión	Descripción
hive-metastore-server	1.0.0-amzn-0	Service para acceder al metaalmacén de Hive, un repositorio semántico que almacena metadatos para SQL en operaciones de Hadoop.
hive-server	1.0.0-amzn-0	Servicio para aceptar consultas de Hive como solicitudes web.
mahout-client	0.10.0	Biblioteca para machine learning.
mysql-server	5.5	Servidor de bases de datos MySQL.
pig-client	0.14.0-amzn-0	Cliente de línea de comando de Pig.
spark-client	1.4.1	Clientes de línea de comando de Spark.
spark-history-server	1.4.1	IU web para la visualiza ción de eventos registrados durante la vida útil de una aplicación Spark completada.
spark-on-yarn	1.4.1	Motor de ejecución en memoria para YARN.
spark-yarn-slave	1.4.1	Bibliotecas de Apache Spark necesarias para esclavos de YARN.

Clasificaciones de configuración de la versión 4.0.0

Las clasificaciones de configuración le permiten personalizar las aplicaciones. Suelen corresponder a un archivo XML de configuración para la aplicación como, por ejemplo, `hive-site.xml`. Para obtener más información, consulte Configuración de aplicaciones.

Clasificaciones de emr-4.0.0

Clasificaciones	Descripción
capacity-scheduler	Cambiar los valores en el archivo capacity-scheduler.xml de Hadoop.
core-site	Cambiar los valores en el archivo core-site.xml de Hadoop.
emrfs-site	Cambiar la configuración de EMRFS.
hadoop-env	Cambiar los valores en el entorno de Hadoop para todos los componentes de Hadoop.
hadoop-log4j	Cambiar los valores en el archivo log4j.properties de Hadoop.
hdfs-site	Cambiar los valores en hdfs-site.xml de HDFS.
hive-env	Cambiar los valores en el entorno de Hive.
hive-exec-log4j	Cambie los valores en el archivo hive-exec-log4j.properties de Hive.
hive-log4j	Cambiar los valores en el archivo hive-log4.properties de Hive.
hive-site	Cambiar los valores en el archivo hive-site.xml de Hive.
httpfs-env	Cambiar los valores en el entorno de HTTPFS.
httpfs-site	Cambiar los valores en el archivo httpfs-site.xml de Hadoop.

Clasificaciones	Descripción
mapred-env	Cambie los valores en el entorno de la MapReduce aplicación.
mapred-site	Cambia los valores en el archivo mapred-site.xml de la MapReduce aplicación.
pig-properties	Cambiar los valores en el archivo pig.properties de Pig.
pig-log4j	Cambiar los valores en el archivo log4j.properties de Pig.
spark	Configuración de Amazon EMR mantenida para Apache Spark.
spark-defaults	Cambiar los valores en el archivo spark-defaults.conf de Spark.
spark-env	Cambiar los valores en el entorno de Spark.
spark-log4j	Cambiar los valores en el archivo log4j.properties de Spark.
yarn-env	Cambiar los valores en el entorno de YARN.
yarn-site	Cambiar los valores en el archivo yarn-site.xml de YARN.

Versiones 2.x y 3.x de la AMI de Amazon EMR

> **Note**
>
> AWS está actualizando la configuración de TLS para todos los puntos de conexión de la AWS API a una versión mínima de TLS 1.2. Las versiones 3.10 y anteriores de Amazon EMR solo admiten conexiones TLS 1.0/1.1. Después del 4 de diciembre de 2023, no podrá crear clústeres con Amazon EMR 3.10 y versiones anteriores.

Si utiliza Amazon EMR 3.10 o una versión anterior, se recomienda que pruebe y migre inmediatamente sus cargas de trabajo a la versión más reciente de Amazon EMR. Para más información, consulte el Blog seguridad de AWS.

Las versiones 2.x y 3.x de Amazon EMR, denominadas versiones de la AMI, están disponibles para las soluciones preexistentes que las necesitan por motivos de compatibilidad. No es recomendable que cree nuevos clústeres o nuevas soluciones con estas versiones. Carecen de las características de las versiones más recientes e incluyen paquetes de aplicaciones obsoletos.

Se recomienda que cree las soluciones con la versión de lanzamiento de Amazon EMR más reciente.

El ámbito de las diferencias entre las versiones de las series 2.x y 3.x y las versiones recientes de Amazon EMR es importante. Estas diferencias abarcan desde cómo crear y configurar un clúster a los puertos y la estructura de directorios de aplicaciones en el clúster.

En esta sección se intentan explicar las diferencias más importantes de Amazon EMR, así como las diferencias de configuración y administración de aplicaciones específicas. No se trata de una lista completa. Si crea y utiliza clústeres de las series 2.x o 3.x, puede encontrar diferencias no descritas en esta sección.

Temas

- Creación de un clúster con versiones anteriores de la AMI de Amazon EMR
- Instalación de aplicaciones con versiones anteriores de la AMI de Amazon EMR
- Personalización de la configuración de clústeres y aplicaciones con versiones anteriores de la AMI de Amazon EMR
- Detalles específicos de las aplicaciones de Hive para versiones anteriores de la AMI de Amazon EMR
- HBase detalles específicos de las aplicaciones para versiones anteriores de la AMI de Amazon EMR
- Detalles específicos de las aplicaciones de Pig para versiones anteriores de la AMI de Amazon EMR
- Detalles específicos de las aplicaciones Spark con versiones anteriores de AMI de Amazon EMR
- Diferencias de la DistCp utilidad S3 con las versiones anteriores de la AMI de Amazon EMR

Creación de un clúster con versiones anteriores de la AMI de Amazon EMR

Las versiones de lanzamiento 2.x y 3.x de Amazon EMR reciben el nombre de versiones de la AMI. Con la versión 4.0.0 y posteriores de Amazon EMR, a las versiones se les referencia por versión de lanzamiento, mediante una etiqueta de versión como `emr-5.11.0`. Este cambio es más evidente cuando se crea un clúster mediante o mediante programación. AWS CLI

Cuando utilice la AWS CLI para crear un clúster mediante una versión de lanzamiento de AMI, utilice la `--ami-version` opción, por ejemplo,`--ami-version 3.11.0`. Muchas opciones, características y aplicaciones introducidas en la versión 4.0.0 y posteriores de Amazon EMR no están disponibles cuando se especifica una `--ami-version`. Para obtener más información, consulte create-cluster en la Referencia de comandos de la AWS CLI .

El siguiente AWS CLI comando de ejemplo lanza un clúster mediante una versión de AMI.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,\
InstanceType=m3.xlarge InstanceGroupType=CORE,InstanceCount=2,\
InstanceType=m3.xlarge --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-
actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Cuando se crea un clúster mediante programación, todas las versiones de Amazon EMR utilizan la acción `RunJobFlowRequest` de la API de EMR. El siguiente código Java de ejemplo crea un clúster con la versión de AMI 3.11.0.

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("AmiVersion Cluster")
   .withAmiVersion("3.11.0")
   .withInstances(new JobFlowInstancesConfig()
```

```
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

La siguiente llamada `RunJobFlowRequest` utiliza una etiqueta de lanzamiento en su lugar:

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("ReleaseLabel Cluster")
   .withReleaseLabel("emr-7.9.0")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

Configuración del tamaño del clúster

Cuando se ejecuta el clúster, Hadoop determina el número de tareas de mapeador y reductor necesarias para procesar los datos. Los clústeres más grandes deben tener más tareas para un mejor uso de recursos y un tiempo de procesamiento más corto. Normalmente, un clúster de EMR sigue teniendo el mismo tamaño durante todo el clúster; define el número de tareas al crear el clúster. Al cambiar el tamaño de un clúster en ejecución, puede variar el procesamiento durante la ejecución del clúster. Por lo tanto, en lugar de utilizar un número de tareas fijo, puede variar el número de tareas durante la vida útil del clúster. Existen dos opciones de configuración para ayudar a definir el número ideal de tareas:

- `mapred.map.tasksperslot`
- `mapred.reduce.tasksperslot`

Puede establecer ambas opciones en el archivo `mapred-conf.xml`. Al enviar un trabajo al clúster, el cliente del trabajo comprueba el número total actual de ranuras de Map y Reduce disponibles en todo el clúster. El cliente de trabajo, a continuación, utiliza las siguientes ecuaciones para definir el número de tareas:

- `mapred.map.tasks` = `mapred.map.tasksperslot` * ranuras de Map en el clúster
- `mapred.reduce.tasks` = `mapred.reduce.tasksperslot` * ranuras de Reduce en el clúster

El cliente de trabajo solo lee el parámetro `tasksperslot` si el número de tareas no está configurado. Puede cambiar el número de tareas en cualquier momento, ya sea para todos los clústeres a través de una acción de arranque o individualmente por trabajo añadiendo un paso para cambiar la configuración.

Amazon EMR resiste los errores de los nodos de tarea y continúa la ejecución del clúster incluso si un nodo de tarea deja de estar disponible. Amazon EMR aprovisiona automáticamente nodos de tareas adicionales para sustituir los que no funcionen.

Puede tener un número diferente de nodos de tareas para cada paso del clúster. También puede añadir un paso a un clúster en ejecución para modificar el número de nodos de tareas. Dado que existe la garantía de que todos los pasos se ejecutarán en secuencia de forma predeterminada, puede especificar el número de nodos de tareas en ejecución para cualquier paso.

Instalación de aplicaciones con versiones anteriores de la AMI de Amazon EMR

Cuando se usa una versión de AMI, las aplicaciones se instalan de varias maneras, incluyendo el `NewSupportedProducts` parámetro de la RunJobFlowacción, las acciones de arranque y la acción Step.

Personalización de la configuración de clústeres y aplicaciones con versiones anteriores de la AMI de Amazon EMR

En la versión 4.0.0 de Amazon EMR se introdujo un método simplificado de configurar aplicaciones mediante clasificaciones de configuración. Para obtener más información, consulte Configuración de aplicaciones. Cuando se utiliza una versión de AMI, las aplicaciones se configuran mediante acciones de arranque y pasando argumentos. Por ejemplo, las acciones de arranque `configure-hadoop` y `configure-daemons` configuran propiedades del entorno específicas de Hadoop y YARN como `--namenode-heap-size`. En versiones más recientes, estas se configuran con las clasificaciones de configuración `hadoop-env` y `yarn-env`. Para ver las acciones de arranque que realizan configuraciones comunes, consulte el emr-bootstrap-actions repositorio en Github.

En las siguientes tablas, se asignan acciones de arranque a clasificaciones de configuración de las versiones de Amazon EMR más recientes.

Hadoop

Nombre de archivo de aplicación afectado	Acción de arranque de versión de la AMI	Clasificación de configuración
`core-site.xml`	`configure-hadoop -c`	`core-site`
`log4j.properties`	`configure-hadoop -l`	`hadoop-log4j`
`hdfs-site.xml`	`configure-hadoop -s`	`hdfs-site`
n/a	n/a	`hdfs-encryption-zones`
`mapred-site.xml`	`configure-hadoop -m`	`mapred-site`
`yarn-site.xml`	`configure-hadoop -y`	`yarn-site`
`httpfs-site.xml`	`configure-hadoop -t`	`httpfs-site`
`capacity-scheduler.xml`	`configure-hadoop -z`	`capacity-scheduler`
`yarn-env.sh`	`configure-daemons --resourcemanager-opts`	`yarn-env`

Hive

Nombre de archivo de aplicación afectado	Acción de arranque de versión de la AMI	Clasificación de configuración
`hive-env.sh`	n/a	`hive-env`
`hive-site.xml`	`hive-script --install -hive-site ${MY_HIVE_SITE_FILE}`	`hive-site`
`hive-exec-log4j.properties`	n/a	`hive-exec-log4j`

Nombre de archivo de aplicación afectado	Acción de arranque de versión de la AMI	Clasificación de configuración
`hive-log4j.properties`	n/a	`hive-log4j`

EMRFS

Nombre de archivo de aplicación afectado	Acción de arranque de versión de la AMI	Clasificación de configuración
`emrfs-site.xml`	`configure-hadoop -e`	`emrfs-site`
n/a	`s3get -s s3://custom-provider.jar -d /usr/share/aws/emr/auxlib/`	`emrfs-site` (con el nuevo valor `fs.s3.cse.encryptionMaterialsProvider.uri`)

Para ver una lista de todas las clasificaciones, consulte Configuración de aplicaciones.

Variables de entorno de la aplicación

Cuando se utiliza una versión de AMI, se usa un script `hadoop-user-env.sh` junto con la acción de arranque `configure-daemons` para configurar el entorno de Hadoop. El script incluye las siguientes acciones:

```
#!/bin/bash
export HADOOP_USER_CLASSPATH_FIRST=true;
echo "HADOOP_CLASSPATH=/path/to/my.jar" >> /home/hadoop/conf/hadoop-user-env.sh
```

En la versión 4.x de Amazon EMR, se hace lo mismo con la clasificación de configuración `hadoop-env`, tal y como se muestra en el ejemplo siguiente:

```
[
      {
         "Classification":"hadoop-env",
         "Properties":{
```

```
        },
        "Configurations":[
           {
              "Classification":"export",
              "Properties":{
                 "HADOOP_USER_CLASSPATH_FIRST":"true",
                 "HADOOP_CLASSPATH":"/path/to/my.jar"
              }
           }
        ]
     }
  ]
```

Otro ejemplo consiste en usar `configure-daemons` y pasar `--namenode-heap-size=2048` y `--namenode-opts=-XX:GCTimeRatio=19`, que es equivalente a las siguientes clasificaciones de configuración.

```
[
     {
        "Classification":"hadoop-env",
        "Properties":{

        },
        "Configurations":[
           {
              "Classification":"export",
              "Properties":{
                 "HADOOP_DATANODE_HEAPSIZE":  "2048",
           "HADOOP_NAMENODE_OPTS":  "-XX:GCTimeRatio=19"
              }
           }
        ]
     }
  ]
```

Otras variables de entorno de aplicaciones ya no están definidas en `/home/hadoop/.bashrc`. En su lugar, se definen principalmente en archivos `/etc/default` por componente o aplicación, como por ejemplo `/etc/default/hadoop`. Los scripts envoltorios `/usr/bin/` instalados por la aplicación también RPMs pueden establecer variables de entorno adicionales antes de incluir el script bin propiamente dicho.

Puertos de servicio

Cuando se utiliza una versión de AMI, algunos servicios utilizan puertos personalizados.

Cambios en la configuración del puerto

Opción	AMI versión 3.x	Opción predeterminada de código abierto
fs.default.name	hdfs://emrDeterminedIP:9000	predeterminado (hdfs: //:8020) *emrDeterminedIP*
dfs.datanode.address	0.0.0.0:9200	valor predeterminado (0.0.0.0:50010)
dfs.datanode.http.address	0.0.0.0:9102	valor predeterminado (0.0.0.0:50075)
dfs.datanode.https.address	0.0.0.0:9402	valor predeterminado (0.0.0.0:50475)
dfs.datanode.ipc.address	0.0.0.0:9201	valor predeterminado (0.0.0.0:50020)
dfs.http.address	0.0.0.0:9101	valor predeterminado (0.0.0.0:50070)
dfs.https.address	0.0.0.0:9202	valor predeterminado (0.0.0.0:50470)
dfs.secondary.http.address	0.0.0.0:9104	valor predeterminado (0.0.0.0:50090)
yarn.nodemanager.address	0.0.0.0:9103	valor predeterminado (${yarn.nodemanager.hostname}:0)
yarn.nodemanager.localizer.address	0.0.0.0:9033	valor predeterminado (${yarn.nodemanager.hostname}:8040)

Opción	AMI versión 3.x	Opción predeterminada de código abierto
yarn.nodemanager.webapp.address	0.0.0.0:9035	valor predeterminado (${yarn.nodemanager.hostname}:8042)
yarn.resourcemanager.address	*emrDeterminedIP* predeterminado (hdfs: //:8020) ----sep----:9022	valor predeterminado (${yarn.resourcemanager.hostname}:8032)
yarn.resourcemanager.admin.address	*emrDeterminedIP* :9022 ----Sept----:9025	valor predeterminado (${yarn.resourcemanager.hostname}:8033)
yarn.resourcemanager.resource-tracker.address	*emrDeterminedIP* :9025 ----SEP----:9023	valor predeterminado (${yarn.resourcemanager.hostname}:8031)
yarn.resourcemanager.scheduler.address	*emrDeterminedIP* :9023 ----Sept----:9024	valor predeterminado (${yarn.resourcemanager.hostname}:8030)
yarn.resourcemanager.webapp.address	0.0.0.0:9026	valor predeterminado (${yarn.resourcemanager.hostname}:8088)
yarn.web-proxy.address	*emrDeterminedIP* :9024 ----Sept----:9046	valor predeterminado (sin valor)
yarn.resourcemanager.hostname	0.0.0.0 (valor predeterminado)	*emrDeterminedIP*

> **Note**
>
> *emrDeterminedIP*Es una dirección IP generada por Amazon EMR.

Usuarios

Cuando se utiliza una versión de AMI, el usuario `hadoop` ejecuta todos los procesos y es propietario de todos los archivos. En la versión 4.0.0 y posteriores de Amazon EMR, los usuarios existen en el nivel de aplicación y de componente.

Secuencia de instalación, artefactos instalados y ubicaciones de archivos de registro

Cuando se utiliza una versión de AMI, los artefactos de aplicación y sus directorios de configuración se instalan en el directorio `/home/hadoop/`*`application`*. Por ejemplo, si instaló Hive, el directorio sería `/home/hadoop/hive`. En la versión 4.0.0 y posteriores de Amazon EMR, los artefactos de la aplicación se instalan en el directorio `/usr/lib/`*`application`*. Cuando se utiliza una versión de AMI, los archivos log se encuentran en diversos lugares. En la tabla siguiente se indican las ubicaciones.

Cambios en las ubicaciones de registro en Amazon S3

Daemon o aplicación	Ubicación del directorio
instance-state	nodo/ /instance-state/ *`instance-id`*
hadoop-hdfs-namenode	*`instance-id`* daemons//.log hadoop-hadoop-namenode
hadoop-hdfs-datanode	daemons/*`instance-id`* /hadoop-hadoop-datanode.log
hadoop-yarn () ResourceManager	daemons/*`instance-id`* /yarn-hadoop-resourcemanager
hadoop-yarn (servidor proxy)	daemones//*`instance-id`* yarn-hadoop-proxyserver
mapred-historyserver	daemones//*`instance-id`*
httpfs	daemons/ /httpfs.log *`instance-id`*
hive-server	nodo/ /hive-server/hive-server.log *`instance-id`*
hive-metastore	nodo/ /apps/hive.log *`instance-id`*

Daemon o aplicación	Ubicación del directorio
Hive CLI	nodo/ /apps/hive.log *instance-id*
Registros de contenedor y de usuario de aplicaciones YARN	task-attempts/
Mahout	N/A
Pig	N/A
spark-historyserver	N/D
archivos de historial de tareas mapreduce	jobs/

Command Runner

Cuando se utiliza una versión de AMI, muchos scripts o programas, como `/home/hadoop/contrib/streaming/hadoop-streaming.jar`, no se colocan en el entorno de ruta de inicio de sesión del shell, por lo que debe especificar la ruta completa cuando utilice un archivo jar como command-runner.jar o script-runner.jar para ejecutar los scripts. `command-runner.jar` está ubicado en la AMI y, por tanto, no hay necesidad de conocer un URI completo como ocurría con `script-runner.jar`.

Factor de replicación

El factor de replicación le permite configurar cuándo iniciar una JVM de Hadoop. Puede iniciar una nueva JVM de Hadoop para cada tarea, lo que proporciona un mejor aislamiento de las tareas, o puede compartirlas JVMs entre tareas, lo que reduce la sobrecarga del marco. Si está procesando muchos archivos pequeños, es razonable reutilizar el JVM muchas veces para amortizar el costo de startup. Sin embargo, si cada tarea tarda mucho tiempo o procesa una gran cantidad de datos, es posible que desee no reutilizar la JVM para garantizar que se libere toda la memoria para las tareas siguientes. Cuando se utiliza una versión de AMI, puede personalizar el factor de replicación mediante la acción de arranque `configure-hadoop` para establecer la propiedad `mapred.job.reuse.jvm.num.tasks`.

El siguiente ejemplo muestra cómo configurar el factor de reutilización de JVM para el uso infinito de la JVM.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m3.xlarge \
InstanceGroupType=CORE,InstanceCount=2,InstanceType=m3.xlarge \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Detalles específicos de las aplicaciones de Hive para versiones anteriores de la AMI de Amazon EMR

Archivos de registro

Cuando se utilizan las versiones 2.x y 3.x de la AMI de Amazon EMR, los registros de Hive se almacenan en `/mnt/var/log/apps/`. Para poder admitir las versiones simultáneas de Hive, la versión de Hive que ejecuta determina el nombre de archivo de registro, tal y como se muestra en la siguiente tabla.

Versión de Hive	Nombre de archivo de registro
0.13.1	`hive.log` **Note** A partir de esta versión, Amazon EMR utiliza un nombre de archivo sin versión, `hive.log`. Las versiones secundarias comparten la misma ubicación de registro que la versión principal.
0.11.0	`hive_0110.log`

Versión de Hive	Nombre de archivo de registro
	Note Las versiones menores de Hive 0.11.0, tales como 0.11.0.1, comparten la misma ubicación de archivo de registro que Hive 0.11.0.
0.8.1	`hive_081.log` **Note** Las versiones menores de Hive 0.8.1, tales como Hive 0.8.1.1, comparten la misma ubicación de archivo de registro que Hive 0.8.1.
0.7.1	`hive_07_1.log` **Note** Las versiones menores de Hive 0.7.1, tales como Hive 0.7.1.3 y Hive 0.7.1.4, comparten la misma ubicación de archivo de registro que Hive 0.7.1.
0.7	`hive_07.log`
0,5	`hive_05.log`
0.4	`hive.log`

Funcionalidad de entrada de división

Para implementar la funcionalidad de entrada de división utilizando versiones de Hive anteriores a la 0.13.1 (versiones de la AMI de Amazon EMR anteriores a la 3.11.0), utilice lo siguiente:

```
hive> set hive.input.format=org.apache.hadoop.hive.ql.io.HiveCombineSplitsInputFormat;
```

```
hive> set mapred.min.split.size=100000000;
```

Esta funcionalidad se dejó de utilizar con Hive 0.13.1. Para obtener la misma funcionalidad de formato de entrada de división en la versión 3.11.0 de la AMI de Amazon EMR, utilice lo siguiente:

```
set hive.hadoop.supports.splittable.combineinputformat=true;
```

Puertos de servicio de Thrift

Thrift es un marco de trabajo de RPC que define un formato de serialización binario compacto utilizado para conservar estructuras de datos para su análisis posterior. Normalmente, Hive configura el servidor para operar en los siguientes puertos.

Versión de Hive	Número de puerto
Hive 0.13.1	10000
Hive 0.11.0	10004
Hive 0.8.1	10003
Hive 0.7.1	10002
Hive 0.7	10001
Hive 0.5	10000

Para obtener más información sobre los servicios de thrift, consulte http://wiki.apache.org/thrift/.

Uso de Hive para recuperar particiones

Amazon EMR incluye una instrucción en el lenguaje de consultas de Hive que recupera las particiones de una tabla de los datos de tabla ubicados en Amazon S3. El ejemplo siguiente lo muestra.

```
CREATE EXTERNAL TABLE (json string) raw_impression
PARTITIONED BY (dt string)
LOCATION 's3://elastic-mapreduce/samples/hive-ads/tables/impressions';
ALTER TABLE logs RECOVER PARTITIONS;
```

Los datos y los directorios de la partición deben estar en la ubicación especificada en la definición de tabla y deben nombrarse de acuerdo con la convención de Hive: por ejemplo, `dt=2009-01-01`.

> **Note**
>
> Después de Hive 0.13.1 esta capacidad se admite de forma nativa `msck repair` *`table`* y, por tanto, no se admite `recover partitions`. Para obtener más información, consulte https://cwiki.apache. org/confluence/display/Hive/LanguageManual+DDL.

Pasar una variable de Hive a un script

Para pasar una variable en un paso de Hive utilizando la AWS CLI, escriba el comando siguiente, *`myKey`* sustituya por el nombre del par de EC2 claves y *`amzn-s3-demo-bucket`* sustituya por el nombre del bucket. En este ejemplo, `SAMPLE` es un valor de variable precedido por el modificador `-d`. Esta variable se define en el script de Hive como: `${SAMPLE}`.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.9 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,\
Args=[-f,s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,\
INPUT=s3://elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/
hive-ads/output/,\
-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

Especificación de la ubicación de un metaalmacén externo

El siguiente procedimiento muestra cómo omitir los valores de configuración predeterminados para la ubicación del metaalmacén de Hive e iniciar un clúster con la ubicación de metaalmacén reconfigurada.

Para crear un metaalmacén situado fuera del clúster de EMR

1. Cree una base de datos MySQL o Aurora mediante Amazon RDS.

   Para obtener información acerca de cómo crear una base de datos de Amazon RDS, consulte Introducción a Amazon RDS.

2. Modifique sus grupos de seguridad para permitir las conexiones JDBC entre su base de datos y el grupo de seguridad ElasticMapReduce-Master.

   Para más información acerca de cómo modificar los grupos de seguridad para el acceso, consulte Grupos de seguridad de Amazon RDS en la Guía del usuario de Amazon RDS.

3. Ajuste los valores de configuración de JDBC en `hive-site.xml`:

   a. Cree un archivo de configuración `hive-site.xml` que contenga lo siguiente:

```
<configuration>
  <property>
    <name>javax.jdo.option.ConnectionURL</name>
    <value>jdbc:mariadb://hostname:3306/hive?createDatabaseIfNotExist=true</
value>
    <description>JDBC connect string for a JDBC metastore</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionUserName</name>
    <value>hive</value>
    <description>Username to use against metastore database</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionPassword</name>
    <value>password</value>
    <description>Password to use against metastore database</description>
  </property>
</configuration>
```

*hostname*es la dirección de DNS de la instancia de Amazon RDS que ejecuta la base de datos. *username*y *password* son las credenciales de su base de datos. Para más información sobre la conexión a instancias de base de datos de MySQL y Aurora, consulte Conexión a la instancia de base de datos que ejecuta el motor de base de datos de MySQL y Conexión a un clúster de base de datos de Aurora en la Guía del usuario de Amazon RDS.

Amazon EMR instala los controladores de JDBC.

> **Note**
>
> La propiedad value no debe contener espacios ni saltos de línea. Debe aparecer todo en una única línea.

b. Guarde el archivo `hive-site.xml` en una ubicación de Amazon S3 como, por ejemplo, `s3://amzn-s3-demo-bucket/hive-site.xml`.

4. Cree un clúster especificando la ubicación en Amazon S3 del archivo `hive-site.xml` personalizado.

El siguiente comando de ejemplo muestra un AWS CLI comando que hace esto.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.10 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--bootstrap-actions Name="Install Hive Site Configuration",\
Path="s3://region.elasticmapreduce/libs/hive/hive-script",\
Args=["--base-path","s3://elasticmapreduce/libs/hive","--install-hive-site",\
"--hive-site=s3://amzn-s3-demo-bucket/hive-site.xml","--hive-versions","latest"]
```

Conectarse a Hive mediante JDBC

Para conectarse a Hive a través de JDBC es necesario que descargue el controlador JDBC e instale un cliente SQL. El siguiente ejemplo muestra el uso de SQL Workbench/J para conectarse a Hive a través de JDBC.

Para descargar controladores de JDBC

1. Descargue y extraiga los controladores adecuados para las versiones de Hive a las que desea acceder. La versión de Hive difiere en función de la AMI que elija al crear un clúster de Amazon EMR.

   - Controladores JDBC Hive 0.13.1: JDBC_1.0.4.1004.zip https://amazon-odbc-jdbc-drivers.s3.amazonaws.com/public/ AmazonHive
   - Controladores JDBC de Hive 0.11.0: https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.11.0
   - Controladores JDBC de Hive 0.8.1: https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.8.1

2. Instale SQL Workbench/J. Para obtener más información, consulte Installing and starting SQL Workbench/J en el manual del usuario de SQL Workbench/J.
3. Cree un túnel de SSH en el nodo principal del clúster. El puerto de conexión varía en función de la versión de Hive. En las tablas siguientes se proporcionan comandos de ejemplo para usuarios `ssh` de Linux y comandos de PuTTY para usuarios de Windows.

   Comandos SSH de Linux

Versión de Hive	Comando
0.13.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10000:localhost:10000 hadoop@` *`master-public-dns-name`*
0.11.0	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10004:localhost:10004 hadoop@` *`master-public-dns-name`*
0.8.1	`ssh -o ServerAliveInterval=10 -i` *`path-to-key-file`* `-N -L 10003:localhost:10003 hadoop@` *`master-public-dns-name`*

Versión de Hive	Comando
0.7.1	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10002:localhost:10002 hadoop@ master-public-dns-name`
0.7	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10001:localhost:10001 hadoop@ master-public-dns-name`
0,5	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000 hadoop@ master-public-dns-name`

Configuración del túnel PuTTY para Windows

Versión de Hive	Ajustes de túnel
0.13.1	Puerto de origen: 10000 Destino: 10000 *master-public-dns-name*
0.11.0	Puerto de origen: 10004 Destino: 10004 *master-public-dns-name*
0.8.1	Puerto de origen: 10003 Destino: 10003 *master-public-dns-name*

4. Añadir el controlador JDBC a SQL Workbench.

   a. En el cuadro de diálogo Select Connection Profile (Seleccionar perfil de conexión), elija Manage Drivers (Administrar controladores).

   b. Elija el icono Create a new entry (Crear entrada nueva) (página en blanco).

   c. En el campo Name (Nombre), escriba **`Hive JDBC`**.

   d. En Library (Biblioteca), haga clic en el icono Select the JAR file(s) (Seleccionar los archivos JAR).

   e. Seleccione los archivos JAR que se muestran en la siguiente tabla.

Versión de controlador de Hive	Archivos JAR que añadir

0.13.1	`hive_metastore.jar`
	`hive_service.jar`
	`HiveJDBC3.jar`
	`libfb303-0.9.0.jar`
	`libthrift-0.9.0.jar`
	`log4j-1.2.14.jar`
	`ql.jar`
	`slf4j-api-1.5.8.jar`
	`slf4j-log4j12-1.5.8.jar`
	`TCLIServiceClient.jar`

0.11.0	`hadoop-core-1.0.3.jar`
	`hive-exec-0.11.0.jar`
	`hive-jdbc-0.11.0.jar`
	`hive-metastore-0.11.0.jar`
	`hive-service-0.11.0.jar`
	`libfb303-0.9.0.jar`
	`commons-logging-1.0.4.jar`
	`slf4j-api-1.6.1.jar`

0.8.1	`hadoop-core-0.20.205.jar`
	`hive-exec-0.8.1.jar`
	`hive-jdbc-0.8.1.jar`
	`hive-metastore-0.8.1.jar`
	`hive-service-0.8.1.jar`
	`libfb303-0.7.0.jar`
	`libthrift-0.7.0.jar`
	`log4j-1.2.15.jar`
	`slf4j-api-1.6.1.jar`
	`slf4j-log4j12-1.6.1.jar`

Versión de controlador de Hive	Archivos JAR que añadir

0.7.1	`hadoop-0.20-core.jar`
	`hive-exec-0.7.1.jar`
	`hive-jdbc-0.7.1.jar`
	`hive-metastore-0.7.1.jar`
	`hive-service-0.7.1.jar`
	`libfb303.jar`
	`commons-logging-1.0.4.jar`
	`slf4j-api-1.6.1.jar`
	`slf4j-log4j12-1.6.1.jar`

0.7	`hadoop-0.20-core.jar`
	`hive-exec-0.7.0.jar`
	`hive-jdbc-0.7.0.jar`
	`hive-metastore-0.7.0.jar`
	`hive-service-0.7.0.jar`
	`libfb303.jar`
	`commons-logging-1.0.4.jar`
	`slf4j-api-1.5.6.jar`
	`slf4j-log4j12-1.5.6.jar`

0,5	`hadoop-0.20-core.jar`
	`hive-exec-0.5.0.jar`
	`hive-jdbc-0.5.0.jar`
	`hive-metastore-0.5.0.jar`
	`hive-service-0.5.0.jar`
	`libfb303.jar`
	`log4j-1.2.15.jar`
	`commons-logging-1.0.4.jar`

f. En el cuadro de diálogo Please select one driver (Seleccione un controlador), seleccione un controlador en función de la siguiente tabla y haga clic en OK (Aceptar).

Versión de Hive	Nombre de clase del controlador
0.13.1	`com.amazon.hive.jdbc3.HS2Driver`
0.11.0	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.8.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0,5	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`

5. Cuando vuelva al cuadro de diálogo Select Connection Profile (Seleccionar perfil de conexión), compruebe que el campo Driver (Controlador) esté definido en Hive JDBC (JDBC de Hive) y proporcione la cadena de conexión de JDBC en el campo URL de acuerdo con la siguiente tabla.

Versión de Hive	Cadena de conexión de JDBC
0.13.1	`jdbc:hive2://localhost:10000/default`
0.11.0	`jdbc:hive://localhost:10004/default`
0.8.1	`jdbc:hive://localhost:10003/default`

Si el clúster utiliza la versión 3.3.1 o posterior de la AMI, en el cuadro de diálogo Select Connection Profile (Seleccionar perfil de conexión), escriba **hadoop** en el campo Username (Nombre de usuario).

HBase detalles específicos de las aplicaciones para versiones anteriores de la AMI de Amazon EMR

Versiones compatibles HBase

HBase versión	Versión de AMI	AWS CLI parámetros de configuración	HBase detalles de la versión

0.94.18	3.1.0 y versiones posteriores	`--ami-version 3.1`
		`--ami-version 3.2`
		`--ami-version 3.3`
		`--applications Name=HBase`
			• Correcciones de errores y mejoras.

0.94.7	3.0-3.0.4	`--ami-version 3.0` `--applications Name=HBase`	
0.92	2.2 y versiones posteriores	`--ami-version 2.2 or later` `--applications Name=HBase`	

HBase requisitos previos de clúster

Un clúster creado con las versiones 2.x y 3.x de la AMI de Amazon EMR debe cumplir los requisitos siguientes para. HBase

- El AWS CLI (opcional): para interactuar con él HBase mediante la línea de comandos, descargue e instale la versión más reciente de. AWS CLI Para obtener más información, consulte Installing the AWS Command Line Interface en la Guía del usuario de AWS Command Line Interface .
- Al menos dos instancias (opcional): el nodo maestro del clúster ejecuta el servidor HBase maestro y ZooKeeper, y los nodos de tareas ejecutan los servidores de la HBase región. Para obtener el máximo rendimiento, HBase los clústeres deben ejecutarse en al menos dos EC2 instancias, pero puede ejecutarse HBase en un único nodo para fines de evaluación.
- Clúster de ejecución prolongada: HBase solo se ejecuta en clústeres de ejecución prolongada. La CLI y la consola de Amazon EMR crean clústeres de ejecución prolongada de manera predeterminada.
- Un conjunto de par de EC2 claves de Amazon (recomendado): para utilizar el protocolo de red Secure Shell (SSH) para conectar con el nodo maestro y ejecutar HBase comandos de intérprete de comandos de Amazon EC2 al crear el clúster.
- Las versiones de AMI y Hadoop correctas (HBase los clústeres) actualmente solo se admiten en Hadoop 20.205 o versiones posteriores.
- Ganglia (opcional): para supervisar las métricas de HBase rendimiento, instale Ganglia al crear el clúster.
- Un bucket de Amazon S3 para registros (opcional): los registros HBase están disponibles en el nodo maestro. Si desea que estos registros se copien en Amazon S3, especifique un bucket de S3 para que reciba archivos de registro al crear el clúster.

Crear un clúster con HBase

En la siguiente tabla se muestran las opciones disponibles cuando se utiliza la consola para crear un clúster HBase con una versión de AMI de Amazon EMR.

Campo	Acción
Restore from backup (Restaurar desde copia de seguridad)	Especifique si desea precargar el HBase clúster con datos almacenados en Amazon S3.
Backup location (Ubicación	Especifique el URI donde reside en Amazon S3 la copia de seguridad desde la que realizará la restauración.

Campo	Acción
de copia de seguridad)	
Backup version (Versión de copia de seguridad)	Opcionalmente, especifique el nombre de la versión de la copia de seguridad en Backup Location (Ubicación de copia de seguridad) que desea utilizar. Si deja este campo en blanco, Amazon EMR utilizará la última copia de seguridad en Ubicación de copia de seguridad para rellenar el nuevo clúster. HBase
Schedule Regular Backups (Programa r copia de seguridads periódicas)	Especifique si desea programar copias de seguridad incrementales automátic as. La primera copia de seguridad es una copia de seguridad completa para crear una referencia para futuros copias de seguridad incrementales.
Consistent backup (Copia de seguridad consistente)	Especifique si las copias de seguridad deben ser consistentes. Una copia de seguridad coherente es la que pone en pausa las operaciones de escritura durante la etapa de copia de seguridad inicial, con sincronización entre nodos. Cualquier operación de escritura que se haya puesto en pausa de este modo se coloca en una cola y se reanuda cuando se completa la sincronización.
Backup frequency (Frecuencia de copia de seguridad)	El número de copias de seguridad programadas days/hours/minutes entre ellas.
Backup location (Ubicación de copia de seguridad)	El URI de Amazon S3 donde se almacenan las copias de seguridad. La ubicación de las copias de seguridad de cada HBase clúster debe ser diferente para garantizar que las copias de seguridad diferenciales sigan siendo correctas.

Campo	Acción
Backup start time (Hora de inicio de copia de seguridad)	Especifica cuándo debe realizarse la primera copia de seguridad. Puede definirlo en `now`, lo que hace que la primera copia de seguridad comience en cuanto el clúster se ejecute, o escriba una fecha y hora con formato ISO. Por ejemplo, 2012-06-15T20:00 Z establecería la hora de inicio en el 15 de junio de 2012 a las 8:00 p. m. UTC.

El siguiente AWS CLI comando de ejemplo lanza un clúster con HBase y otras aplicaciones:

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
               --applications Name=Hue Name=Hive Name=Pig Name=HBase \
               --use-default-roles --ec2-attributes KeyName=myKey \
               --instance-type c1.xlarge --instance-count 3 --termination-protected
```

Una vez realizada la conexión entre el Hive y HBase los clústeres (como se muestra en el procedimiento anterior), puede acceder a los datos almacenados en el HBase clúster creando una tabla externa en Hive.

En el siguiente ejemplo, cuando se ejecuta desde el símbolo del sistema de Hive, se crea una tabla externa que hace referencia a los datos almacenados en una HBase tabla denominada`inputTable`. A continuación, puede hacer referencia `inputTable` a las instrucciones de Hive para consultar y modificar los datos almacenados en el HBase clúster.

> **Note**
>
> El siguiente ejemplo utiliza protobuf-java-2.4.0a.jar en la AMI 2.3.3, pero debe modificar el ejemplo de adaptarlo a su versión. Para comprobar la versión de Protocol Buffers JAR que tiene, ejecute el comando en el símbolo del sistema de Hive: `! ls /home/hadoop/lib;`.

```
add jar lib/emr-metrics-1.0.jar ;
               add jar lib/protobuf-java-2.4.0a.jar ;

               set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com ;

               create external table inputTable (key string, value string)
                    stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
                     with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
                     tblproperties ("hbase.table.name" = "t1");

               select count(*) from inputTable ;
```

Personalización HBase de la configuración

Aunque la configuración predeterminada debería funcionar con la mayoría de aplicaciones, tiene la flexibilidad de modificar sus opciones HBase de configuración. Para ello, ejecute uno de los dos scripts de acciones de arranque:

- configure-hbase-daemons—Configura propiedades de los daemons maestros, regionserver y de zookeeper. Estas propiedades incluyen el tamaño del montón y opciones que transferir a la Máquina virtual Java (JVM) cuando se inicia el HBase daemon. Puede definir estas propiedades como argumentos en la acción de arranque. Esta acción de arranque modifica el archivo de configuración/home/hadoop/conf/hbase-user-env.sh del clúster. HBase
- configure-hbase: configura opciones HBase específicas del sitio tales como el puerto al que debe vincular el HBase maestro y el número máximo de veces que el cliente de la CLI debe volver a intentar una acción. Puede definirlas one-by-one como argumentos en la acción de arranque o puede especificar la ubicación de un archivo de configuración XML en Amazon S3. Esta acción de arranque modifica el archivo de configuración the /home/hadoop/conf/hbase -site.xml en el clúster. HBase

> **Note**
>
> Estos scripts, al igual que otras acciones de arranque, solo se pueden ejecutar cuando se crea el clúster; no se pueden usar para cambiar la configuración de un HBase clúster que se esté ejecutando actualmente.

Al ejecutar las acciones configure-hbase o configure-hbase-daemonsbootstrap, los valores que especifique anulan los valores predeterminados. Cualquier valor que no haya definido explícitamente toma los valores predeterminados.

La configuración HBase con estas acciones de arranque es análoga al uso de acciones de arranque en Amazon EMR para configurar opciones de Hadoop y propiedades de daemon de Hadoop. La diferencia es que no tiene opciones de memoria por proceso. HBase En su lugar, las opciones de memoria se configuran mediante el `--`*`daemon`*`-opts` argumento, donde *`daemon`* se sustituye por el nombre del daemon a configurar.

Configurar HBase daemons

Amazon EMR proporciona una acción de arranque que puede utilizar para cambiar la configuración de HBase daemons, que *`region`* es la región en la que desea lanzar el clúster. `s3://`*`region`*`.elasticmapreduce/bootstrap-actions/configure-hbase-daemons` HBase

Para configurar los HBase daemons mediante el AWS CLI, añada la acción de `configure-hbase-daemons` arranque al lanzar el clúster para configurar uno o más daemons. HBase Puede definir las siguientes propiedades:

Propiedad	Descripción
`hbase-master-opts`	Opciones que controlan cómo la JVM ejecuta el daemon maestro. Si están definidas, anulan las variables HBASE_MASTER_OPTS predeterminadas.
`regionserver-opts`	Opciones que controlan cómo la JVM ejecuta el daemon de servidor de región. Si están definidas, anulan las variables HBASE_REGIONSERVER_OPTS predeterm inadas.
`zookeeper-opts`	Opciones que controlan cómo la JVM ejecuta el daemon zookeeper. Si están definidas, anulan las variables HBASE_ZOOKEEPER_OPTS predeterminadas.

Para obtener más información sobre estas opciones, consulte el archivo hbase-env.sh en la documentación. HBase

En el siguiente ejemplo se muestra una acción de arranque para configurar valores para `zookeeper-opts` y `hbase-master-opts`.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-daemons,\
Args=["--hbase-zookeeper-opts=-Xmx1024m -XX:GCTimeRatio=19","--hbase-master-opts=-Xmx2048m","--hbase-regionserver-opts=-Xmx4096m"]
```

Configurar ajustes de HBase sitio

Amazon EMR proporciona una acción de arranque`s3://elasticmapreduce/bootstrap-actions/configure-hbase`, que puede utilizar para cambiar la configuración. HBase Puede definir valores de configuración one-by-one, como argumentos en la acción de arranque o puede especificar la ubicación de un archivo de configuración XML en Amazon S3. Establecer los valores de configuración one-by-one es útil si solo necesita establecer algunos ajustes de configuración. Definirlos utilizando un archivo XML resulta útil si tiene que hacer muchos cambios o si desea guardar las opciones de configuración que reutilizar.

> **Note**
>
> Puede agregar un prefijo de región al nombre del bucket de Amazon S3`s3://region.elasticmapreduce/bootstrap-actions/configure-hbase`, como dónde *`region`* está la región donde se va a lanzar el clúster HBase .

Esta acción de arranque modifica el archivo de `/home/hadoop/conf/hbase-site.xml` configuración del clúster. HBase La acción de arranque solo se puede ejecutar cuando se lanza el HBase clúster.

Para obtener más información sobre los ajustes del HBase sitio que se pueden configurar, consulte Configuración predeterminada en la HBase documentación.

Establezca la acción de `configure-hbase` arranque al lanzar el HBase clúster y especifique los valores `hbase-site.xml` que desee cambiar.

Para especificar la configuración individual HBase del sitio mediante el AWS CLI

- Para cambiar la `hbase.hregion.max.filesize` configuración, escriba el comando siguiente y *myKey* sustituya por el nombre de su par de EC2 claves de Amazon.

  > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
hbase,Args=["-s","hbase.hregion.max.filesize=52428800"]
```

Para especificar la configuración del HBase sitio con un archivo XML, utilice la AWS CLI

1. Cree una versión personalizada de `hbase-site.xml`. Su archivo personalizado debe ser XML válido. Para reducir la probabilidad de introducir errores, empiece por la copia predeterminada de`hbase-site.xml`, que se encuentra en el nodo HBase maestro de Amazon EMR en`/home/hadoop/conf/hbase-site.xml`, y edite una copia de dicho archivo en lugar de crearlo desde cero. Puede proporcionar un nombre nuevo al nuevo archivo o dejarlo como `hbase-site.xml`.
2. Cargue el archivo `hbase-site.xml` personalizado en un bucket de Amazon S3. Debe tener permisos establecidos de modo que la AWS cuenta de que lanza el clúster pueda acceder al archivo. Si la AWS cuenta de que lanza el clúster también es propietaria del bucket de Amazon S3, tendrá acceso.
3. Defina la acción bootstrap configure-hbase al lanzar el HBase clúster e incluya la ubicación del archivo personalizado. `hbase-site.xml` El siguiente ejemplo establece los valores de

configuración del HBase sitio en los especificados en el archivo. `s3://amzn-s3-demo-bucket/my-hbase-site.xml` Escriba el comando siguiente, *myKey* sustituya el nombre de su par de EC2 claves y *amzn-s3-demo-bucket* sustituya por el nombre de su bucket de Amazon S3.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
        --applications Name=Hue Name=Hive Name=Pig Name=HBase \
        --use-default-roles --ec2-attributes KeyName=myKey \
        --instance-type c1.xlarge --instance-count 3 --termination-protected \
        --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
hbase,Args=["--site-config-file","s3://amzn-s3-demo-bucket/config.xml"]
```

Si especifica más de una opción para personalizar la HBase operación, debe anteponer a cada par clave-valor un modificador de `-s` opción, como se muestra en el siguiente ejemplo:

```
      --bootstrap-actions s3://elasticmapreduce/bootstrap-actions/configure-
hbase,Args=["-s","zookeeper.session.timeout=60000"]
```

Con el proxy definido y la conexión SSH abierta, puede ver la HBase IU cargando http: //:60010/ master-status en una ventana de navegador*master-public-dns-name*, donde se encuentra *master-public-dns-name* la dirección DNS pública del nodo maestro del clúster. HBase

Puede consultar los HBase registros actuales utilizando SSH para conectar al nodo maestro y acceder al `mnt/var/log/hbase` directorio. Estos registros no estarán disponibles una vez que el clúster se haya terminado a menos que hubiera habilitado el registro en Amazon S3 cuando se lanzó el clúster.

Hacer una copia de seguridad y restaurar HBase

Amazon EMR le brinda la posibilidad de hacer copias de seguridad de HBase los datos en Amazon S3, ya sea manualmente o mediante una programación automatizada. Puede realizar tanto copias de seguridad completas como incrementales. Una vez que haya hecho una copia de seguridad de HBase los datos, puede restaurarla en un clúster. HBase Puede restaurar un HBase clúster que se esté ejecutando actualmente o lanzar un clúster nuevo rellenado previamente con los datos de los que se ha hecho una copia de seguridad.

Durante el proceso de copia de seguridad, HBase sigue ejecutando los comandos de escritura. Aunque esto garantiza que el clúster permanezca disponible durante la copia de seguridad, existe el riesgo de inconsistencia entre los datos cuya copia de seguridad se realiza y las operaciones de escritura que se ejecutan en paralelo. Para entender las inconsistencias que puedan surgir, hay que tener en cuenta que HBase distribuye las operaciones de escritura entre los nodos de su clúster. Si una operación de escritura se produce después de sondear un nodo determinado, dichos datos no se incluyen en el archivo de copia de seguridad. Es posible que incluso descubras que las escrituras anteriores en el HBase clúster (enviadas a un nodo que ya ha sido sondeado) podrían no estar en el archivo de respaldo, mientras que las escrituras posteriores (enviadas a un nodo antes de ser sondeado) sí están incluidas.

Si se requiere una copia de seguridad coherente, debe pausar las escrituras HBase durante la parte inicial del proceso de copia de seguridad, es decir, la sincronización entre los nodos. Puede hacerlo especificando el parámetro `--consistent` al solicitar una copia de seguridad. Con este parámetro, las escrituras durante este período se ponen en cola y se ejecutan en cuanto se completa la sincronización. También puede programar copias de seguridad recurrentes, que resuelven cualquier inconsistencia a lo largo del tiempo, ya que los datos que faltan en un paso de copia de seguridad se incluyen en el siguiente paso de copia de seguridad.

Al hacer una copia de seguridad de HBase los datos, debe especificar un directorio de copia de seguridad diferente para cada clúster. Una forma sencilla de hacerlo consiste en utilizar el identificador del clúster como parte de la ruta especificada en el directorio de copia de seguridad. Por ejemplo, `s3://amzn-s3-demo-bucket/backups/j-3AEXXXXXX16F2`. Esto garantiza que cualquier copia de seguridad incremental futura haga referencia al HBase clúster correcto.

Cuando esté listo para eliminar los archivos de copia de seguridad antiguos que ya no son necesarios, le recomendamos que primero haga una copia de seguridad completa de sus HBase datos. Esto garantiza que se conserven todos los datos y ofrece una línea de base para futuras copias de seguridad incrementales. Una vez realizado la copia de seguridad completa, puede

acceder a la ubicación de copia de seguridad y eliminar manualmente los archivos de copia de seguridad antiguos.

El proceso HBase de copia de seguridad utiliza S3 DistCp para la operación de copia, que presenta ciertas limitaciones en cuanto al espacio de almacenamiento temporal de archivos.

Hacer una copia de seguridad y restaurar HBase utilizando la consola

La consola permite lanzar un nuevo clúster y rellenarlo con los datos de una HBase copia de seguridad anterior. También le permite programar copias de seguridad incrementales periódicas de HBase los datos. Otra funcionalidad adicional de copia de seguridad y restauración, como la posibilidad de restaurar datos en un clúster que ya está en ejecución, realizar copias de seguridad manuales y programar copias de seguridad completos automatizadas, está disponible utilizando la CLI.

Para rellenar un nuevo clúster con HBase datos archivados mediante la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Create cluster.
3. En la sección Configuración de software, para Aplicaciones adicionales, elija Configurar HBasey agregar.
4. En el cuadro de diálogo Add Application (Añadir aplicación), marque Restore From Backup (Restaurar a partir de una copia de seguridad).
5. En Backup Location, especifique la ubicación de la copia de seguridad para cargarla en el nuevo HBase clúster. Debe ser una URL de Amazon S3 de la forma `s3://amzn-s3-demo-bucket/backups/`.
6. Puede definir un valor para Backup Version (Versión de copia de seguridad) para especificar el nombre de la versión de copia de seguridad que desea cargar. Si no define un valor para la Versión de copia de seguridad, Amazon EMR carga la copia de seguridad más reciente en la ubicación especificada.
7. Elija Add (Añadir) y continúe con la creación del clúster especificando las demás opciones que desee.

Para programar copias de seguridad automatizadas de HBase los datos mediante la consola

1. En la sección Configuración del software, para Aplicaciones adicionales, elija Configurar HBasey añadir.
2. Elija Schedule Regular Backups (Programar copias de seguridad periódicas).
3. Especifique si las copias de seguridad deben ser consistentes. Una copia de seguridad coherente es la que pone en pausa las operaciones de escritura durante la etapa de copia de seguridad inicial, con sincronización entre nodos. Cualquier operación de escritura que se haya puesto en pausa de este modo se coloca en una cola y se reanuda cuando se completa la sincronización.
4. Defina la frecuencia con la que se deben realizar las copias de seguridad introduciendo un número para Backup Frequency (Frecuencia de copia de seguridad) y eligiendo Days (Días), Hours (Horas) o Minutes (Minutos). La primera copia de seguridad automatizada que se ejecuta es una copia de seguridad completa; posteriormente, Amazon EMR guarda las copias de seguridad incrementales en función de la programación que especifique.
5. Especifique la ubicación en Amazon S3 donde deben almacenarse las copias de seguridad. Se debe realizar una copia de seguridad de cada HBase clúster en una ubicación independiente en Amazon S3 para garantizar que las copias de seguridad incrementales se calculen correctamente.
6. Especifique cuándo se debe realizar la primera copia de seguridad definiendo un valor para Backup Start Time (Hora de inicio de copia de seguridad). Puede definirlo en `now`, lo que hace que la primera copia de seguridad comience en cuanto el clúster se ejecute, o escriba una fecha y hora con formato ISO. Por ejemplo, 2013-09-26T20:00 Z establece la hora de inicio en el 26 de septiembre de 2013 a las 8:00 p. m. UTC.
7. Elija Agregar.
8. Continúe con la creación del clúster con otras opciones como desee.

Supervise HBase con CloudWatch

Amazon EMR notifica tres métricas CloudWatch que puede utilizar para supervisar sus HBase copias de seguridad. Estas métricas se envían a CloudWatch intervalos de cinco minutos y se proporcionan sin costo alguno.

Métrica	Descripción
`HBaseBackupFailed`	Si la última copia de seguridad produjo un error. Está establecido en 0 de forma predeterminada y se actualiza a 1 si el intento de copia de seguridad anterior produjo un error. Esta métrica solo se informa para HBase los clústeres. Caso de uso: monitorizar HBase copias de seguridad Unidades: recuento
`HBaseMostRecentBackupDuration`	La cantidad de tiempo que tarda en completarse la copia de seguridad anterior. Esta métrica se establece independientemente de si la última copia de seguridad se realizó correctamente o produjo un error. Mientras la copia de seguridad está en curso, esta métrica devuelve el número de minutos tras el inicio de la copia de seguridad. Esta métrica solo se informa para HBase los clústeres. Caso de uso: monitorizar HBase copias de seguridad Unidades: minutos
`HBaseTimeSinceLastSuccessfulBackup`	El número de minutos transcurridos desde que se inició la última HBase copia de seguridad correcta en el clúster. Esta métrica solo se indica para HBase los clústeres. Caso de uso: monitorizar HBase copias de seguridad Unidades: minutos

Configure Ganglia para HBase

Se configura Ganglia para HBase usar la acción configure-hbase-for-gangliabootstrap. Esta acción de arranque se configura para publicar métricas en HBase Ganglia.

Debe configurar HBase Ganglia al lanzar el clúster; los informes de Ganglia no se pueden agregar a un clúster en ejecución.

Ganglia también almacena los archivos de registro en el servidor en `/mnt/var/log/ganglia/rrds`. Si configura su clúster para mantener los archivos de registro en un bucket de Amazon S3, los archivos de registro de Ganglia también se mantienen ahí.

Para lanzar un clúster con Ganglia como función HBase, utilice la acción de configure-hbase-for-gangliaarranque, tal y como se muestra en el siguiente ejemplo.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase Name=Ganglia \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-for-ganglia
```

Una vez que el clúster se haya lanzado con Ganglia configurado, puede acceder a los informes y gráficos de Ganglia utilizando la interfaz gráfica que se ejecuta en el nodo principal.

Detalles específicos de las aplicaciones de Pig para versiones anteriores de la AMI de Amazon EMR

Versiones de Pig compatibles

La versión de Pig que puede agregar a su clúster depende de la versión de la AMI de Amazon EMR y de la versión de Hadoop que utilice. En la tabla siguiente se muestra las versiones de la AMI y las versiones de Hadoop que son compatibles con las distintas versiones de Pig. Le recomendamos que utilice la última versión disponible de Pig para aprovechar las mejoras de rendimiento y las nuevas funcionalidades.

Cuando utilizas la API para instalar Pig, se utiliza la versión predeterminada, a menos que especifiques `--pig-versions` como argumento el paso que carga Pig en el clúster durante la llamada a. RunJobFlow

Versión de Pig	Versión de AMI	Parámetros de configuración	Detalles de la versión de Pig

0.12.0
Notas de la versión
Documentación
	3.1.0 y versiones posteriores	`--ami-version 3.1`
		`--ami-version 3.2`
		`--ami-version 3.3`
			Añade compatibilidad para lo siguiente:
			• Transmisión UDFs sin implement aciones de JVM
			• Operadores ASSERT e IN
			• Expresión CASE
			• AvroStorage como función integrada de Pig.
			• ParquetLoader y ParquetStorer como funciones integrada s
			• BigInteger y BigDecimal tipos

0.11.1.1 Notas de la versión Documentación	2.2 y versiones posteriores	`--pig-versions 0.11.1.1` `--ami-version 2.2`	Mejora el rendimien to del comando LOAD PigStorage si la entrada reside en Amazon S3.
0.11.1 Notas de la versión Documentación	2.2 y versiones posteriores	`--pig-versions 0.11.1` `--ami-version 2.2`	Agrega compatibi lidad con JDK 7, Hadoop 2, funciones de Groovy definidas por el usuario,

Versión de Pig	Versión de AMI	Parámetros de configuración	Detalles de la versión de Pig
			SchemaTuple optimización, nuevos operadores, etc. Para más información, consulte el Registro de cambios de Pig 0.11.1.
0.9.2.2 Notas de la versión Documentación	2.2 y versiones posteriores	`--pig-versions 0.9.2.2` `--ami-version 2.2`	Añade compatibilidad para Hadoop 1.0.3.
0.9.2.1 Notas de la versión Documentación	2.2 y versiones posteriores	`--pig-versions 0.9.2.1` `--ami-version 2.2`	Añade soporte para MapR.
0.9.2 Notas de la versión Documentación	2.2 y versiones posteriores	`--pig-versions 0.9.2` `--ami-version 2.2`	Incluye varias mejoras de rendimiento y correcciones de errores. Para obtener información completa sobre los cambios de Pig 0.9.2, consulte el registro de cambios de Pig 0.9.2.
0.9.1 Notas de la versión Documentación	2.0	`--pig-versions 0.9.1` `--ami-version 2.0`	

Versión de Pig	Versión de AMI	Parámetros de configuración	Detalles de la versión de Pig

0.6
 	
Notas de la versión
	1.0	`--pig-versions 0.6`
		
		`--ami-version 1.0`	

0.3
 	
Notas de la versión
	1.0	`--pig-versions 0.3`
		
		`--ami-version 1.0`	

Detalles de la versión de Pig

Amazon EMR admite determinadas versiones de Pig que podrían tener aplicados parches de Amazon EMR adicionales. Puede configurar la versión de Pig que desea ejecutar en clústeres de Amazon EMR. Para obtener más información acerca de cómo hacerlo, consulte Apache Pig. En las siguientes secciones se describen diferentes versiones de Pig y los parches aplicados a las versiones cargadas en Amazon EMR.

Parches de Pig

En esta sección se describen los parches personalizados aplicados a versiones de Pig disponibles con Amazon EMR.

Parches de Pig 0.11.1.1

La versión de Amazon EMR de Pig 0.11.1.1 es una versión de mantenimiento que mejora el rendimiento del comando LOAD PigStorage si la entrada reside en Amazon S3.

Parches de Pig 0.11.1

La versión de Amazon EMR de Pig 0.11.1 contiene todas las actualizaciones facilitadas por la Apache Software Foundation y los parches acumulativos de Amazon EMR de la versión 0.9.2.2 de Pig. Sin embargo, no hay parches nuevos específicos de Amazon EMR en Pig 0.11.1.

Parches de Pig 0.9.2

Apache Pig 0.9.2 es una versión de mantenimiento de Pig. El equipo de Amazon EMR ha aplicado los siguientes parches a la versión de Amazon EMR de Pig 0.9.2.

Parche	Descripción
PIG-1429	Añadir el tipo de datos booleano a Pig como tipo de datos de primer nivel. Para obtener más información, consulte https://issues.apache. org/jira/browse/PIG-1429. Estado: confirmado Corregido en la versión de Apache Pig: 0.10
PIG-1824	Admite módulos de importación en Jython UDF. Para obtener más información, consulte https://issues.apache. org/jira/browse/PIG-1824. Estado: confirmado Corregido en la versión de Apache Pig: 0.10
PIG-2010	Paquete registrado JARs en la memoria caché distribuida. Para obtener más información, consulte https://issues.apache. org/jira/browse/PIG-2010. Estado: confirmado Corregido en la versión de Apache Pig: 0.11
PIG-2456	Añadir un archivo ~/.pigbootup donde el usuario puede especificar instrucciones de Pig predeterminadas. Para obtener más información, consulte https://issues.apache. org/jira/browse/PIG-2456. Estado: confirmado Corregido en la versión de Apache Pig: 0.11

Parche	Descripción
PIG-2623	Support utilizar rutas de Amazon S3 para registrarse UDFs. Para obtener más información, consulte https://issues.apache.org/jira/browse/PIG-2623. Estado: confirmado Corregido en la versión de Apache Pig: 0.10, 0.11

Parches de Pig 0.9.1

El equipo de Amazon EMR ha aplicado los siguientes parches a la versión de Amazon EMR de Pig 0.9.1.

Parche	Descripción
Admite archivos de JAR y scripts de Pig en dfs	Agregue compatibilidad para ejecutar scripts y registrar archivos de JAR almacenados en HDFS, Amazon S3 u otros sistemas de archivos distribuidos. Para obtener más información, consulte https://issues.apache. org/jira/browse/PIG-1505. Estado: confirmado Corregido en la versión de Apache Pig: 0.8.0
Soporta varios sistemas de archivos en Pig	Añade soporte para scripts de Pig para leer datos desde un sistema de archivos y escribirlos en otro. Para obtener más información, consulte https://issues.apache. org/jira/browse/PIG-1564. Estado: no confirmado Corregido en la versión de Apache Pig: n/d
Agregue la fecha, hora y cadena de Piggybank UDFs	Agregue la fecha, la hora y la cadena para admitir scripts Pig personalizados. UDFs Para obtener más información, consulte https://issues.apache. org/jira/browse/PIG-1565.

Parche	Descripción
	Estado: no confirmado Corregido en la versión de Apache Pig: n/d

Clústeres de Pig interactivos y por lotes

Amazon EMR le permite ejecutar scripts de Pig en dos modos:

- Interactivo
- Lote

Al lanzar un clúster de larga duración mediante la consola o la AWS CLI, puede conectarse ssh al nodo maestro como usuario de Hadoop y utilizar el shell de Grunt para desarrollar y ejecutar sus scripts de Pig de forma interactiva. El uso de Pig de forma interactiva le permite revisar el script de Pig con mayor facilidad que el modo por lotes. Después de revisar correctamente el script de Pig en modo interactivo, puede cargar el script en Amazon S3 y utilizar el modo por lotes para ejecutar el script en la fase de producción. También puede enviar comandos de Pig de forma interactiva en un clúster en ejecución para analizar y transformar los datos según sea necesario.

En el modo de lotes, se carga el script de Pig en Amazon S3 y, a continuación, se envía el trabajo al clúster como un paso. Los pasos de Pig se pueden enviar a un clúster en ejecución prolongada o a un clúster transitorio.

Detalles específicos de las aplicaciones Spark con versiones anteriores de AMI de Amazon EMR

Usar Spark interactivamente o en modo de lote

Amazon EMR le permite ejecutar aplicaciones de Spark en dos modos:

- Interactivo
- Lote

Cuando lanzas un clúster de larga duración con la consola o la consola AWS CLI, puedes conectarte mediante SSH al nodo maestro como usuario de Hadoop y usar el shell de Spark para desarrollar

y ejecutar tus aplicaciones de Spark de forma interactiva. El uso de Spark de forma interactiva le permite crear prototipos o probar aplicaciones de Spark con mayor facilidad en un entorno de lotes. Después de revisar correctamente la aplicación Spark en modo interactivo, puede colocar la aplicación JAR o el programa Python en el sistema de archivos local en el nodo maestro del clúster en Amazon S3. A continuación, puede enviar la aplicación como un flujo de trabajo por lotes.

En el modo de lotes, cargue su script de Spark en Amazon S3 o el sistema de archivos del nodo maestro local y, a continuación, envíe el trabajo al clúster como un paso. Los pasos de Spark se pueden enviar a un clúster en ejecución prolongada o a un clúster transitorio.

Creación de un clúster con Spark instalado

Para lanzar un clúster con Spark instalado utilizando la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Create cluster.
3. En Software Configuration (Configuración de software), elija la versión de la AMI que necesite.
4. Para Applications to be installed (Aplicaciones que hay que instalar), elija Spark en la lista y, a continuación, elija Configure and add (Configurar y añadir).
5. Añada argumentos para cambiar la configuración de Spark como desee. Para obtener más información, consulte Configurar Spark. Elija Add (Agregar).
6. Seleccione las demás opciones que necesite y, a continuación, elija Create cluster (Crear clúster).

El siguiente ejemplo muestra cómo crear un clúster con Spark mediante Java:

```
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
SupportedProductConfig sparkConfig = new SupportedProductConfig()
   .withName("Spark");

RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("Spark Cluster")
   .withAmiVersion("3.11.0")
   .withNewSupportedProducts(sparkConfig)
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyName")
```

```
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge")
   );
RunJobFlowResult result = emr.runJobFlow(request);
```

Configurar Spark

Para configurar Spark al crear un clúster, ejecuta la acción de arranque que se encuentra en el repositorio de Github. awslabs/emr-bootstrap-actions/spark Para ver los argumentos que acepta la acción de arranque, consulte README en ese repositorio. La acción de arranque configura propiedades en el archivo `$SPARK_CONF_DIR/spark-defaults.conf`. Para obtener más información acerca de la configuración, consulte el tema Spark Configuration en la documentación de Spark. Puede sustituir "latest" en la URL siguiente por el número de versión de Spark que va a instalar, por ejemplo, `2.2.0` http://spark.apache.org/docs/latest/configuration.html.

También es posible configurar Spark dinámicamente en el momento de enviar la aplicación. Existe un valor para maximizar automáticamente la asignación de recursos para un ejecutor utilizando el archivo de configuración `spark`. Para obtener más información, consulte Anulación de las opciones de configuración predeterminadas de Spark.

Cambio de la configuración predeterminada de Spark

El siguiente ejemplo muestra cómo crear un clúster con `spark.executor.memory` establecido en 2G mediante la AWS CLI.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Spark cluster" --ami-version 3.11.0 \
--applications Name=Spark, Args=[-d,spark.executor.memory=2G] --ec2-attributes
 KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 --use-default-roles
```

Enviar trabajo a Spark

Para enviar trabajos a un clúster, utilice un paso para ejecutar el script `spark-submit` en su clúster de EMR. Agrega el paso mediante el `addJobFlowSteps` método que se muestra en AmazonElasticMapReduceClient:

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
StepFactory stepFactory = new StepFactory();
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

StepConfig sparkStep = new StepConfig()
 .withName("Spark Step")
 .withActionOnFailure("CONTINUE")
 .withHadoopJarStep(stepFactory.newScriptRunnerStep("/home/hadoop/spark/bin/spark-
submit","--class","org.apache.spark.examples.SparkPi","/home/hadoop/spark/lib/spark-
examples-1.3.1-hadoop2.4.0.jar","10"));

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

Anulación de las opciones de configuración predeterminadas de Spark

Es posible que desee anular los valores de configuración predeterminados de Spark por aplicación. Puede hacerlo cuando envíe aplicaciones utilizando un paso, que básicamente transfiere opciones a `spark-submit`. Por ejemplo, es posible que desee cambiar la memoria asignada a un proceso de ejecutor cambiando `spark.executor.memory`. Puede proporcionar el modificador `--executor-memory` con un argumento como el siguiente:

```
/home/hadoop/spark/bin/spark-submit --executor-memory 1g --class
 org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/spark-examples*.jar 10
```

Del mismo modo, puede ajustar `--executor-cores` y `--driver-memory`. En un paso, debería proporcionar los siguientes argumentos al paso:

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/
spark-examples*.jar 10
```

También puede ajustar la configuración que puede no tener un modificador integrado utilizando la opción `--conf`. Para obtener más información sobre otras opciones que se pueden ajustar, consulte el tema Dynamically loading Spark properties en la documentación de Apache Spark.

Diferencias de la DistCp utilidad S3 con las versiones anteriores de la AMI de Amazon EMR

DistCp Versiones de S3 compatibles con Amazon EMR

Las siguientes DistCp versiones de S3 se admiten en las versiones de la AMI de Amazon EMR. DistCpLas versiones de S3 posteriores a la 1.0.7 se encuentran directamente en los clústeres. Utilice el JAR de `/home/hadoop/lib` para las características más recientes.

Versión	Descripción	Fecha de lanzamiento de la nueva versión
1.0.8	Añade las opciones `--appendToLastFile` , `--requirePreviousManifest` y `--storageClass` .	3 de enero de 2014
1.0.7	Añade la opción `--s3ServerSideEncryption` .	2 de mayo de 2013
1.0.6	Añade la opción `--s3Endpoint` .	6 de agosto de 2012
1.0.5	Mejora la capacidad de especificar qué versión de S3 se va DistCp a ejecutar.	27 de junio de 2012
1.0.4	Mejora la opción `--deleteOnSuccess` .	19 de junio de 2012
1.0.3	Añade soporte para las opciones `--numberFiles` y `--startingIndex` .	12 de junio de 2012
1.0.2	Mejora la nomenclatura de archivos al utilizar grupos.	6 de junio de 2012

Versión	Descripción	Fecha de lanzamiento de la nueva versión
1.0.1	Versión inicial de S3DistCp.	19 de enero de 2012

Agregar un paso de DistCp copia de S3 a un clúster

Para agregar un paso de DistCp copia de S3 a un clúster en ejecución, escriba el comando siguiente, *j-3GYXXXXXX9IOK* sustituya el ID del clúster y *amzn-s3-demo-bucket* sustituya por el nombre del bucket de Amazon S3.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--s3Endpoint,s3-eu-west-1.amazonaws.com",\
"--src,s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node/",\
"--dest,hdfs:///output",\
"--srcPattern,.*[a-zA-Z,]+"]
```

Example Carga de CloudFront registros de Amazon en HDFS

Este ejemplo carga CloudFront registros de Amazon en HDFS agregando un paso a un clúster en ejecución. En el proceso, cambia el formato de compresión de Gzip (el valor predeterminado de CloudFront) a LZO. Esto resulta útil porque los datos comprimidos mediante LZO se pueden dividir en varias asignaciones a medida que se descomprimen, por lo que ya no tendrá que esperar hasta que se complete la compresión, como hace con Gzip. Esto proporciona un mejor rendimiento al analizar los datos con Amazon EMR. Este ejemplo mejora además el rendimiento utilizando la expresión regular especificada en la opción `--groupBy` para combinar todos los registros de una hora dada en un archivo único. Los clústeres de Amazon EMR son más eficaces al procesar

unos pocos archivos grandes comprimidos con LZO que al procesar muchos archivos pequeños comprimidos con Gzip. Para dividir los archivos de LZO, debe indexarlos y utilizar la biblioteca de terceros hadoop-lzo.

Para cargar CloudFront registros de Amazon S3, escriba el comando siguiente, *j-3GYXXXXXX9IOK* sustituya el ID del clúster y sustituya *amzn-s3-demo-bucket* el nombre del bucket de Amazon S3.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--src,s3://amzn-s3-demo-bucket/cf","--dest,hdfs:///local",\
"--groupBy,.*XABCD12345678.([0-9]+-[0-9]+-[0-9]+-[0-9]+).*",\
"--targetSize,128",
"--outputCodec,lzo","--deleteOnSuccess"]
```

Tenga en cuenta el caso en el que el ejemplo anterior se ejecuta sobre los siguientes archivos de registro de CloudFront .

```
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.HLUS3JKx.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.I9CNAZrg.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.YRRwERSA.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.dshVLXFE.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.LpLfuShd.gz
```

S3 DistCp copia, concatena y comprime los archivos en los dos archivos siguientes, donde el nombre del archivo viene determinado por la coincidencia realizada por la expresión regular.

```
hdfs:///local/2012-02-23-01.lzo
hdfs:///local/2012-02-23-02.lzo
```

Novedades

En esta página, se describen los cambios y las funcionalidades disponibles en las versiones más recientes de Amazon EMR 7.x, 6.x y 5.x.

Estas notas de la versión también están disponibles en las páginas Amazon EMR 7.9.0, Amazon EMR 6.15.0 y Amazon EMR 5.36.2, junto con las versiones de la aplicación, las versiones de los componentes y las clasificaciones de configuración disponibles para cada versión.

- Para ver las notas de lanzamiento de versiones anteriores, consulte el Archivo de notas de la versión de Amazon EMR.
- Para recibir actualizaciones cuando haya una nueva versión de Amazon EMR disponible, suscríbase a la fuente RSS de las notas de la versión de Amazon EMR.

> **Note**
>
> Las versiones posteriores de Amazon EMR utilizan AWS Signature Version 4 (SigV4) para autenticar las solicitudes a Amazon S3. Le recomendamos que utilice una versión de Amazon EMR compatible con SigV4 para acceder a nuevos buckets de S3 y evitar la interrupción de sus cargas de trabajo. Para obtener más información y ver una lista de las versiones de Amazon EMR compatibles con SigV4, consulte Amazon EMR y AWS Signature versión 4.

Amazon EMR 7.9.0 (última versión de la serie 7.x)

Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Las siguientes notas de la versión incluyen información sobre la versión 7.9.0 de Amazon EMR.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.
- Compatibilidad con CM-CMK para la interfaz de usuario persistente de aplicaciones: esta función amplía el cifrado SSE aplicado a los registros gestionados en el depósito de servicios de EMR para cumplir con la clave maestra personalizada gestionada por el cliente (CM-CMK). Esta función permite a los clientes gestionar la accesibilidad de los registros una vez finalizados sus clústeres.

Problemas conocidos

- Cuando la opción de registro para el EMR activado EC2 está deshabilitada, los registros no están disponibles una vez finalizado el clúster y la aplicación persistente no UIs contendrá ningún registro.

Cambios, mejoras y problemas resueltos

- Obsolescencia de Oozie: dado que OSS Oozie ha pasado al modo ático, la última versión que incluye Apache Oozie es la versión 7.9 de Amazon EMR. Apache Oozie se eliminará de la versión 7.10 y versiones posteriores de Amazon EMR. Le recomendamos que utilice MWAA como alternativa a Oozie. Los clientes pueden utilizar la utilidad para migrar los flujos oozie-to-airflow-emrde trabajo de Oozie existentes GitHub a MWAA.
- Embalaje de contenedores YARN: a partir de la versión 7.9.0 de Amazon EMR, la política de embalaje de contenedores ya está disponible para el programador de capacidad de YARN, que se basa en la política de ubicación de varios nodos de YARN.
- El archivo hudi-cli.sh: el script hudi-cli.sh quedó obsoleto en la versión 7.9.0 y posteriores de EMR. La versión 7.9.0 y superior de Amazon EMR usa .jar. hudi-cli-bundle
- Clústeres de EMR en una AMI AL2 023 habilitada para FIPS: a partir de EMR 7.9, los clientes pueden lanzar clústeres de EMR en una AMI 023 habilitada para FIPS. AL2 Este cambio actualiza las firmas digitales y los algoritmos de hash de EMR para garantizar que todos los paquetes de EMR cumplan con las políticas criptográficas recomendadas para el FIPS 140-3. Nota: Esto no cubre el estado de la certificación FIPS de EMR AL2 o 023; consulte las preguntas frecuentes sobre la FIPS 023 AL2para obtener la información más reciente sobre el estado de validación de la FIPS 140-3.

- Latencia escalonada: esta versión resuelve un problema por el que la latencia escalonada aumentaba cuando se utilizaba el cifrado en tránsito sin seleccionar explícitamente la aplicación Hadoop para su instalación.
- Lanzamiento de clústeres con volúmenes cifrados con EBS: esta versión corrige un problema por el que una llamada de KMS en volúmenes cifrados con EBS podía provocar un error en el lanzamiento de un clúster de EMR.

-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250527.1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyu Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central), Asia Pacífico (Tokio)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250218.1	6.1.128-136.201. amzn2023	08 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Occidental (Calgary), Asia Pacífico (Malasia), Asia Pacífico (Malasia), México (Central)

Amazon EMR 6.15.0 (última versión de la serie 6.x)

Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Las siguientes notas de la versión incluyen información sobre la versión 6.15.0 de Amazon EMR. Los cambios son respecto a la versión 6.14.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.15.0.

Nuevas características

- Actualizaciones de la aplicación: Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.
- Lanzamientos más rápidos para clústeres de EMR que se ejecutan en EC2: ahora es hasta un 35% más rápido lanzar Amazon EMR en un clúster. EC2 Con esta mejora, la mayoría de los clientes pueden lanzar sus clústeres en 5 minutos o menos.
- CodeWhisperer para EMR Studio: ahora puede utilizar Amazon con CodeWhisperer Amazon EMR Studio para obtener recomendaciones en tiempo real a medida que escribe código. JupyterLab

CodeWhisperer puede completar sus comentarios, terminar líneas de código individuales, hacer line-by-line recomendaciones y generar funciones con formato completo.

- Tiempos de reinicio de tareas más rápidos con Flink: con Amazon EMR 6.15.0 y versiones posteriores, Apache Flink dispone de varios mecanismos nuevos para mejorar el tiempo de reinicio de los trabajos durante las operaciones de recuperación o escalado de tareas. Esto optimiza la velocidad de los gráficos de recuperación y reinicio de la ejecución para mejorar la estabilidad del trabajo.
- Control de acceso detallado y a nivel de tabla para formatos de tabla abierta: con Amazon EMR 6.15.0 y versiones posteriores, cuando ejecuta trabajos de Spark en Amazon EMR EC2 en clústeres que acceden a los datos del catálogo de datos de Glue, AWS Lake Formation puede utilizarlos para aplicar permisos a nivel de tabla, fila, columna y celda en AWS tablas basadas en Hudi, Iceberg o Delta Lake.
- Actualización de Hadoop: Amazon EMR 6.15.0 incluye una actualización de Apache Hadoop a la versión 3.3.6. Hadoop 3.3.6 era la versión más reciente en el momento de la implementación de Amazon EMR 6.15, lanzada por Apache en junio de 2023. Las versiones anteriores de Amazon EMR (6.9.0 a 6.14.x) utilizaban Hadoop 3.3.3.

  La actualización incluye cientos de mejoras y correcciones, además de características que incluyen parámetros de nodos de datos reconfigurables, la opción `DFSAdmin` para iniciar operaciones de reconfiguración masiva en todos los nodos de datos activos y una API vectorizada que permite a los lectores más exigentes especificar varios rangos de lectura. Hadoop 3.3.6 también añade compatibilidad con HDFS APIs y semántica para su registro de escritura anticipada (WAL), de modo que HBase pueda ejecutarse en otras implementaciones de sistemas de almacenamiento. Para obtener más información, consulte los registros de cambios de las versiones 3.3.4, 3.3.5 y 3.3.6 en la documentación de Apache Hadoop.
- Soporte para AWS SDK for Java, versión 2: las aplicaciones de Amazon EMR 6.15.0 pueden usar las versiones 1.12.569 o 2.20.160 AWS del SDK para Java si la aplicación es compatible con la versión 2. El AWS SDK para Java 2.x es una importante reescritura del código base de la versión 1.x. Se basa en Java 8 y agrega varias características solicitadas con frecuencia. Entre estas se incluyen la compatibilidad con operaciones de E/S sin bloqueo y la capacidad de conectar una implementación HTTP diferente en tiempo de ejecución. Para obtener más información, incluida una guía de migración del SDK para Java v1 a la v2, consulte la guía AWS SDK para Java, versión 2.

Cambios, mejoras y problemas resueltos

- Para mejorar sus clústeres de EMR de alta disponibilidad, esta versión permite la conectividad con los daemons de Amazon EMR en un host local que utiliza puntos de conexión IPv6.
- Esta versión permite que TLS 1.2 se comunique con todos los nodos principales del clúster de alta disponibilidad ZooKeeper aprovisionado.
- Esta versión mejora la administración de los archivos de registro de ZooKeeper transacciones que se mantienen en los nodos principales para minimizar los escenarios en los que los archivos de registro se exceden de los límites e interrumpen las operaciones del clúster.
- Esta versión hace que la comunicación entre nodos sea más resiliente para los clústeres de EMR de alta disponibilidad. Esta mejora reduce la posibilidad de que se produzcan errores en las acciones de arranque o en el inicio del clúster.
- En Amazon EMR 6.15.0, Tez presenta configuraciones que puede especificar para abrir de forma asíncrona las divisiones de entrada en una división agrupada de Tez. Esto se traduce en un rendimiento más rápido de las consultas de lectura cuando hay un gran número de divisiones de entrada en una sola división agrupada de Tez. Para obtener más información, consulte Apertura de división asíncrona de Tez.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(España), China (Ningxia), Europa (Zúrich)
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich), Asia Pacífico (Malasia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	13 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

Amazon EMR 5.36.2 (versión más reciente de la serie 5.x)

Las nuevas versiones de Amazon EMR estarán disponibles en diferentes regiones durante un periodo de varios días, comenzando por la primera región en la fecha de lanzamiento inicial. Es posible que la versión de lanzamiento más reciente no esté disponible en su región durante este periodo.

Las siguientes notas de la versión incluyen información sobre la versión 5.36.2 de Amazon EMR. Los cambios son respecto a la versión 5.36.1. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Cambios, mejoras y problemas resueltos

- Esta versión mejora la lógica de reducción vertical del clúster para que Amazon EMR no intente reducir verticalmente los nodos básicos por debajo de la configuración del factor de replicación de HDFS del clúster. Esta mejora cumple con los requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- Esta versión agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se ejecute indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.
- Se corrige un problema en el que las operaciones de reducción de escala del clúster pueden detenerse mientras Amazon EMR desactiva un nodo básico de forma gradual y este pasa a mal estado antes de desactivarse por completo..

- Se mejora la estabilidad de un nodo en un clúster de alta disponibilidad con varios nodos básicos cuando Amazon EMR reinicia un único nodo.
- Optimiza la administración de registros con Amazon EMR que se ejecuta en Amazon. EC2 Como resultado, es posible que observe una pequeña reducción en los costos de almacenamiento de los registros de su clúster.
- Mejora la administración de los archivos de registro de ZooKeeper transacciones que se mantienen en los nodos principales para minimizar los escenarios en los que los archivos de registro se exceden de los límites e interrumpen las operaciones del clúster.
- Corrige un error poco frecuente que podía provocar el fallo de un clúster de alta disponibilidad con varios nodos principales por no poder comunicarse con el Yarn. ResourceManager
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 503.0	4,1343	xxxxxx, 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

Amazon EMR y AWS Signature versión 4

Las versiones de Amazon EMR utilizan la versión 4 (SigV4) de AWS Signature para autenticar las solicitudes a Amazon S3. Los buckets creados en Amazon S3 después del 24 de junio de 2020 no admiten solicitudes firmadas por Signature Version 2 (SigV2). Los buckets creados el 24 de junio de 2020 o antes seguirán siendo compatibles con SigV2. Le recomendamos que migre a una versión de Amazon EMR compatible con SigV4 para acceder a nuevos buckets de S3 y evitar la interrupción de sus cargas de trabajo.

Si utiliza aplicaciones incluidas en Amazon EMR, como Apache Spark, Apache Hive y Presto, no necesita cambiar el código de la aplicación para usar SigV4. Si utiliza aplicaciones personalizadas que no están incluidas en Amazon EMR, es posible que deba actualizar el código para utilizar SigV4. Para más información, consulte Transición de Signature Version 2 a Signature Version 4 en la Guía del usuario de Amazon S3.

Las siguientes versiones de Amazon EMR son compatibles con SigV4: emr-4.7.4, emr-4.8.5, emr-4.9.6, emr-4.10.1, emr-5.1.1, emr-5.2.3, emr-5.3.2, emr-5.4.1, emr-5.5.4, emr-5.6.1, emr-5.7.1, emr-5.8.3, emr-5.9.1, emr-5.10.1, emr-5.11.4, emr-5.12.3, emr-5.13.1, emr-5.14.2, emr-5.15.1, emr-5.16.1, emr-5.17.2, emr-5.18.1, emr-5.19.1, emr-5.20.1, emr-5.21.2, and emr-5.22.0 and higher. Todas las versiones 6.x y 7.x son compatibles con SigV4.

Método para mitigar CVE-2021-44228

> **Note**
>
> En el caso de la versión 6.9.0 y posteriores de Amazon EMR, todos los componentes instalados por Amazon EMR que utilizan bibliotecas Log4j utilizan la versión 2.17.1 o posteriores de Log4j.

Amazon EMR ejecutándose en EC2

El problema que se analiza en CVE-2021-44228 afecta a las versiones principales de Apache Log4j comprendidas entre la 2.0.0 y la 2.14.1 cuando se procesan entradas de orígenes que no son de confianza. Los clústeres de Amazon EMR lanzados con las versiones 5.x de Amazon EMR hasta la 5.34.0 y las versiones 6.x de EMR hasta Amazon EMR 6.5.0 incluyen marcos de código abierto como Apache Hive, Flink, HUDI, Presto y Trino, que utilizan estas versiones de Apache Log4j. Sin

embargo, muchos clientes utilizan los marcos de código abierto instalados en sus clústeres de Amazon EMR para procesar y registrar las entradas de orígenes que no son de confianza.

Se recomienda que aplique la "Solución de acción de arranque de Amazon EMR para Log4j CVE-2021-44228" como se describe en la siguiente sección. Esta solución también aborda CVE-2021-45046.

> **Note**
>
> Los scripts de acción de arranque para Amazon EMR se actualizaron el 7 de septiembre de 2022 para incluir mejoras y correcciones de errores incrementales para Oozie. Si utiliza Oozie, debe aplicar la solución de acción de arranque de Amazon EMR actualizada que se describe en la siguiente sección.

Amazon EMR en EKS

Si utiliza Amazon EMR en EKS con la configuración predeterminada, no se verá afectado por el problema descrito en CVE-2021-44228 y no tendrá que aplicar la solución descrita en la sección Solución de acción de arranque de Amazon EMR para Log4j CVE-2021-44228 y CVE-2021-45046. Para Amazon EMR en EKS, el tiempo de ejecución de Amazon EMR para Spark utiliza la versión 1.2.17 de Apache Log4j. Cuando utilice Amazon EMR en EKS, no debe cambiar la configuración predeterminada del componente `log4j.appender` a `log`.

Solución de acción de arranque de Amazon EMR para Log4j CVE-2021-44228 y CVE-2021-45046

Esta solución proporciona una acción de arranque de Amazon EMR que debe aplicarse en los clústeres de Amazon EMR. Para cada versión de Amazon EMR, encontrará a continuación un enlace a un script de acciones de arranque. Para aplicar esta acción de arranque, debe completar los siguientes pasos:

1. Copie el script correspondiente a su versión de Amazon EMR en un bucket de S3 local en su Cuenta de AWS. Asegúrese de utilizar un script de arranque específico para su versión de Amazon EMR.
2. Configure una acción de arranque para sus clústeres de EMR a fin de ejecutar el script copiado en su bucket de S3 según las instrucciones que se describen en la documentación de EMR. Si

tiene otras acciones de arranque configuradas para sus clústeres de EMR, asegúrese de que este script esté configurado como el primer script de acciones de arranque que se ejecute.

3. Termine los clústeres de EMR existentes y lance nuevos clústeres con el script de acción bootstrap. AWS recomienda probar los scripts de arranque en el entorno de pruebas y validar las aplicaciones antes de aplicarlos al entorno de producción. Si no está utilizando la revisión más reciente para una versión secundaria de EMR (por ejemplo, 6.3.0), debe usar la revisión más reciente (por ejemplo, 6.3.1) y, a continuación, aplicar la solución descrita anteriormente.

CVE-2021-44228 y CVE-2021-45046: scripts de arranque para versiones de Amazon EMR

Número de versión de Amazon EMR	Ubicación del script	Fecha de publicación del script
6.5.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.5.0-v2.sh`	24 de marzo de 2022
6.4.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.4.0-v2.sh`	24 de marzo de 2022
6.3.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.3.1-v2.sh`	24 de marzo de 2022
6.2.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.2.1-v2.sh`	24 de marzo de 2022
6.1.1	`s3://elasticmapreduce/bootstrap-actions/`	14 de diciembre de 2021

Número de versión de Amazon EMR	Ubicación del script	Fecha de publicación del script
	`log4j/patch-log4j-emr-6.1.1-v2.sh`	
6.0.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.0.1-v2.sh`	14 de diciembre de 2021
5,34,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.34.0-v2.sh`	12 de diciembre de 2021
5.33.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.33.1-v2.sh`	12 de diciembre de 2021
5.32,1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.32.1-v2.sh`	13 de diciembre de 2021
5,31.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.31.1-v2.sh`	13 de diciembre de 2021
5.30,2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.30.2-v2.sh`	14 de diciembre de 2021

Número de versión de Amazon EMR	Ubicación del script	Fecha de publicación del script
5.29,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.29.0-v2.sh`	14 de diciembre de 2021
5.28,1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.28.1-v2.sh`	15 de diciembre de 2021
5.27.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.27.1-v2.sh`	15 de diciembre de 2021
5.26.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.26.0-v2.sh`	15 de diciembre de 2021
5.25,0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.25.0-v2.sh`	15 de diciembre de 2021
5.24.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.24.1-v2.sh`	15 de diciembre de 2021

Número de versión de Amazon EMR	Ubicación del script	Fecha de publicación del script
5,23.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.23.1-v2.sh`	15 de diciembre de 2021
5.22.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.22.0-v2.sh`	15 de diciembre de 2021
5.21.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.21.2-v2.sh`	15 de diciembre de 2021
5.20.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.20.1-v2.sh`	15 de diciembre de 2021
5.19,1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.19.1-v2.sh`	15 de diciembre de 2021
5.18,1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.18.1-v2.sh`	15 de diciembre de 2021

Número de versión de Amazon EMR	Ubicación del script	Fecha de publicación del script
5.17.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.17.2-v2.sh`	15 de diciembre de 2021
5.16.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.16.1-v2.sh`	15 de diciembre de 2021
5.15.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.15.1-v2.sh`	15 de diciembre de 2021
5.14.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.14.2-v2.sh`	15 de diciembre de 2021
5.13.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.13.1-v2.sh`	15 de diciembre de 2021
5.12.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.12.3-v2.sh`	15 de diciembre de 2021

Número de versión de Amazon EMR	Ubicación del script	Fecha de publicación del script
5.11.4	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.11.4-v2.sh`	15 de diciembre de 2021
5.10.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.10.1-v2.sh`	15 de diciembre de 2021
5.9.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.9.1-v2.sh`	15 de diciembre de 2021
5.8.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.8.3-v2.sh`	15 de diciembre de 2021
5.7.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.7.1-v2.sh`	15 de diciembre de 2021

Versión de lanzamiento de EMR	Revisión más reciente desde diciembre de 2021
6.3.0	6.3.1
6.2.0	6.2.1
6.1.0	6.1.1

Versión de lanzamiento de EMR	Revisión más reciente desde diciembre de 2021
6.0.0	6.0.1
5.33.0	5.33.1
5,32,0	5.32,1
5,31,0	5,31.1
5.30.0 o 5.30.1	5.30,2
5.28,0	5.28,1
5.27.0	5.27.1
5.24.0	5.24.1
5,23,0	5,23.1
5.21.0 o 5.21.1	5.21.2
5.20.0	5.20.1
5.19,0	5.19,1
5.18,0	5.18,1
5.17.0 o 5.17.1	5.17.2
5.16.0	5.16.1
5.15.0	5.15.1
5.14.0 o 5.14.1	5.14.2
5.13.0	5.13.1
5.12.0, 5.12.1, 5.12.2	5.12.3
5.11.0, 5.11.1, 5.11.2, 5.11.3	5.11.4

Versión de lanzamiento de EMR	Revisión más reciente desde diciembre de 2021
5.9.0	5.9.1
5.8.0, 5.8.1, 5.8.2	5.8.3
5.7.0	5.7.1

Preguntas frecuentes

- ¿Las versiones de EMR anteriores a EMR 5 se ven afectadas por CVE-2021-44228?

  No. Las versiones de EMR anteriores a la versión 5 de EMR utilizan versiones de Log4j anteriores a la 2.0.

- ¿Esta solución también aborda el CVE-2021-45046?

  Sí. Esta solución también aborda el CVE-2021-45046.

- ¿La solución gestiona las aplicaciones personalizadas que instalo en mis clústeres de EMR?

  El script de arranque solo actualiza los archivos JAR que instala EMR. Si instala y ejecuta aplicaciones personalizadas y archivos JAR en sus clústeres de EMR mediante acciones de arranque, como pasos enviados a sus clústeres, con una AMI de Amazon Linux personalizada o por medio de cualquier otro mecanismo, póngase en contacto con el proveedor de la aplicación para determinar si sus aplicaciones personalizadas se ven afectadas por CVE-2021-44228 y determinar la solución adecuada.

- ¿Cómo debo gestionar las imágenes de docker personalizadas con EMR en EKS?

  Si añade aplicaciones personalizadas a Amazon EMR en EKS mediante imágenes de docker personalizadas o envía trabajos a Amazon EMR en archivos de solicitud EKSwith personalizados, póngase en contacto con el proveedor de la aplicación para determinar si sus aplicaciones personalizadas se ven afectadas por el CVE-2021-44228 y determinar la solución adecuada.

- ¿Cómo funciona el script de arranque para mitigar el problema descrito en CVE-2021-44228 y CVE-2021-45046?

  El script de arranque actualiza las instrucciones de inicio de EMR al agregar un nuevo conjunto de instrucciones. Estas nuevas instrucciones eliminan los archivos de JndiLookup clase utilizados

a través de Log4j por todos los marcos de código abierto instalados por EMR. Esto sigue la recomendación publicada por Apache para abordar los problemas de Log4j.

- ¿Hay alguna actualización de EMR que utilice la versión 2.17.1 o posteriores de Log4j?

  Las versiones de EMR 5 hasta la versión 5.34 y las versiones de EMR 6 hasta la versión 6.5 utilizan versiones anteriores de marcos de código abierto que son incompatibles con las versiones más recientes de Log4j. Si continúa utilizando estas versiones, le recomendamos que aplique la acción de arranque para mitigar los problemas descritos en la. CVEs Después de EMR 5 versión 5.34 y EMR 6 versión 6.5, las aplicaciones que utilicen Log4j 1.x y Log4j 2.x se actualizarán para utilizar Log4j 1.2.17 (o superior) y Log4j 2.17.1 (o superior) respectivamente, y no requerirán usar las acciones de arranque mencionadas anteriormente para mitigar los problemas de CVE.

- ¿Las versiones de EMR se ven afectadas por el CVE-2021-45105?

  CVE-2021-45105 no afecta a las aplicaciones instaladas por Amazon EMR con las configuraciones predeterminadas de EMR. De las aplicaciones instaladas por Amazon EMR, solo Apache Hive utiliza Apache Log4j con búsquedas de contexto y no utiliza un diseño de patrones no predeterminado de manera que permita procesar datos de entrada inapropiados.

- ¿Se ve afectado Amazon EMR por alguna de las siguientes divulgaciones de CVE?

  En la siguiente tabla se incluye una lista de CVEs los datos relacionados con Log4j y se indica si cada CVE afecta a Amazon EMR. La información de esta tabla solo se aplica cuando Amazon EMR instala las aplicaciones con las configuraciones predeterminadas.

CVE	Afecta a EMR	Notas
CVE-2022-23302	No	Amazon EMR no configura Log4j JMSSink
CVE-2022-23305	No	Amazon EMR no configura Log4j JDBCAppender
CVE-2022-23307	No	Amazon EMR no configura Log4j Chainsaw
CVE-2020-9493	No	Amazon EMR no configura Log4j Chainsaw

CVE	Afecta a EMR	Notas
CVE-2021-44832	No	Amazon EMR no configura Log4j JDBCAppender con una cadena de conexión JNDI
CVE-2021-4104	No	Amazon EMR no utiliza Log4j JMSAppender
CVE-2020-9488	No	Las aplicaciones que instala Amazon EMR no utilizan Log4j SMTPAppender
CVE-2019-17571	No	Amazon EMR bloquea el acceso público a los clústeres y no los lanza SocketServer
CVE-2019-17531	No	Recomendamos que actualice a la versión más reciente de Amazon EMR. Amazon EMR 5.33.0 y versiones posteriores utilizan jackson-databind 2.6.7.4 o posteriores, y EMR 6.1.0 y versiones posteriores utilizan jackson-databind 2.10.0 o posteriores. Estas versiones de jackson-databind no se ven afectadas por el CVE.

Archivo de notas de la versión de Amazon EMR

Las notas de la versión de todas las versiones de Amazon EMR están disponibles a continuación. Para ver información de la versión completa para cada versión, consulte Versiones de lanzamiento de Amazon EMR 6.x, Versiones de lanzamiento de Amazon EMR 5.x y Versiones de lanzamiento de Amazon EMR 4.x.

Para recibir actualizaciones cuando haya una nueva versión de Amazon EMR disponible, suscríbase a la fuente RSS de las notas de la versión de Amazon EMR.

Versión 6.14.0

Las siguientes notas de la versión incluyen información sobre la versión 6.14.0 de Amazon EMR. Los cambios son respecto a la versión 6.13.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.14.0.

Nuevas características

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- El Escalado administrado de Amazon EMR ya está disponible en la región Asia-Pacífico (Yakarta) `ap-southeast-3` para los clústeres que cree con Amazon EMR 6.14.0 y versiones posteriores.

Cambios, mejoras y problemas resueltos

- La versión 6.14.0 optimiza la administración de registros con Amazon EMR que se ejecuta en Amazon. EC2 Como resultado, es posible que observe una pequeña reducción en los costos de almacenamiento de los registros de su clúster.
- La versión 6.14.0 mejora el flujo de trabajo de escalado para tener en cuenta las diferentes instancias principales que tienen una variación sustancial de tamaño en sus volúmenes de Amazon EBS. Esta mejora se aplica únicamente a los nodos principales; las operaciones de reducción vertical de la escala de los nodos de tareas no se ven afectadas.
- La versión 6.14.0 mejora la forma en que Amazon EMR interactúa con aplicaciones de código abierto, como Apache Hadoop YARN ResourceManager and HDFS NameNode. Esta mejora reduce el riesgo de que se produzcan retrasos operativos relacionados con el escalado de clústeres y mitiga las fallas en el arranque que se producen debido a problemas de conectividad con las aplicaciones de código abierto.
- La versión 6.14.0 optimiza la instalación de las aplicaciones durante el lanzamiento del clúster. Esto reduce la cantidad necesaria de tiempo para que arranque el clúster para determinadas combinaciones de aplicaciones de Amazon EMR.
- La versión 6.14.0 corrige un problema por el que las operaciones de reducción vertical de la escala del clúster se interrumpen cuando un clúster que se ejecuta en una VPC con un dominio personalizado experimenta el reinicio de un nodo de tareas o nodo principal.

- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich), Asia Pacífico (Malasia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	17 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	11 de septiembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

Versión 6.13.0

Las siguientes notas de la versión incluyen información sobre la versión 6.13.0 de Amazon EMR. Los cambios son respecto a la versión 6.12.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.13.0.

Nuevas características

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

Cambios, mejoras y problemas resueltos

- La versión 6.13.0 mejora el daemon de administración de registros de Amazon EMR para garantizar que todos los registros se carguen a una cadencia normal en Amazon S3 cuando se emite un comando de terminación de clúster. Esto facilita que los clústeres se terminen de forma más rápida.
- La versión 6.13.0 mejora las capacidades de administración de registros de Amazon EMR para garantizar la carga coherente y puntual de todos los archivos de registro a Amazon S3. Esto beneficia especialmente a los clústeres de EMR de ejecución prolongada.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

Versión 6.12.0

Las siguientes notas de la versión incluyen información sobre la versión 6.12.0 de Amazon EMR. Los cambios son respecto a la versión 6.11.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.12.0.

Nuevas características

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Las versiones 6.12.0 y posteriores de Amazon EMR admiten la integración de LDAP con Apache Livy, Apache Hive through HiveServer 2 (HS2), Trino, Presto y Hue. También puede instalar Apache Spark y Apache Hadoop en un clúster de EMR que utilice la versión 6.12.0 o posteriores y configurarlos para que utilicen LDAP. Para más información, consulte Uso de servidores de Active Directory o LDAP para la autenticación con Amazon EMR.

Cambios, mejoras y problemas resueltos

- La versión 6.12.0 y posteriores de Amazon EMR proporcionan compatibilidad con el tiempo de ejecución de Java 11 para Flink. Para obtener más información, consulte Configuración de Flink para que se ejecute con Java 11.
- La versión 6.12.0 agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se detenga indefinidamente debido a una única operación de cambio de tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.
- La versión 6.12.0 corrige un problema que provocaba que las operaciones de reducción vertical de los clústeres se interrumpieran cuando un nodo principal que se encontraba en proceso de retirada estable pasara a un estado incorrecto por cualquier motivo antes de retirarse por completo.
- La versión 6.12.0 mejora la lógica de reducción vertical de los clústeres para que el clúster no intente reducir verticalmente los nodos principales por debajo de la configuración de la configuración del factor de replicación de HDFS del clúster. Esto se ajusta a sus requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- La versión 6.12.0 mejora el rendimiento y la eficiencia del servicio de supervisión de estado de Amazon EMR al aumentar la velocidad a la que registra los cambios de estado de las instancias.

Esta mejora reduce la posibilidad de que se degrade el rendimiento de los nodos del clúster que ejecutan varias herramientas de cliente personalizadas o aplicaciones de terceros.

- La versión 6.12.0 mejora el rendimiento del daemon de administración de registros en el clúster para Amazon EMR. Como resultado, hay menos probabilidades de que se degrade el rendimiento con los clústeres de EMR que ejecutan pasos con una alta simultaneidad.
- Con la versión 6.12.0 de Amazon EMR, el daemon de administración de registros se ha actualizado para identificar todos los registros que están en uso activo con controladores de archivos abiertos en el almacenamiento de instancias local y los procesos asociados. Esta actualización garantiza que Amazon EMR elimine correctamente los archivos y recupere espacio de almacenamiento una vez archivados los registros en Amazon S3.
- La versión 6.12.0 incluye una mejora del daemon de administración de registros que elimina los directorios escalonados vacíos y no utilizados del sistema de archivos del clúster local. Un número excesivamente elevado de directorios vacíos puede reducir el rendimiento de los daemons de Amazon EMR y provocar una sobreutilización del disco.
- La versión 6.12.0 permite la rotación de registros para los registros del servidor YARN Timeline. Esto minimiza los escenarios de sobreutilización del disco, especialmente en el caso de clústeres de ejecución prolongada.
- El tamaño del volumen raíz predeterminado ha aumentado a 15 GB en la versión 6.10.0 y posteriores de Amazon EMR. Las versiones anteriores tienen un tamaño de volumen raíz predeterminado de 10 GB.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Oregón) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Canadá Occidental (Calgary),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Europa (España), China (Ningxia), Europa (Zúrich)
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

Versión 6.11.1

Las siguientes notas de la versión incluyen información sobre la versión 6.11.1 de Amazon EMR. Los cambios son respecto a la versión 6.11.0. Para obtener información sobre el cronograma de lanzamiento, consulte el Registro de cambios de la versión 6.11.1.

Cambios, mejoras y problemas resueltos

- Debido a la contención de bloqueo, un nodo puede entrar en un interbloqueo si se agrega o elimina al mismo tiempo que intenta retirarse. Como resultado, el administrador de recursos de Hadoop (YARN) deja de responder y afecta a todos los contenedores entrantes y en ejecución actualmente.
- Esta versión incluye un cambio que permite que los clústeres de alta disponibilidad se recuperen de un estado de error tras el reinicio.
- Esta versión incluye correcciones de seguridad para Hue y. HBase
- Esta versión corrige un problema por el que los clústeres que ejecutan cargas de trabajo en Spark con Amazon EMR podrían recibir silenciosamente resultados incorrectos con `contains`, `startsWith`, `endsWith` y `like`. Este problema se produce cuando utiliza las expresiones en campos particionados que tienen metadatos en Hive3 Metastore Server (HMS) de Amazon EMR.
- Esta versión corrige un problema con la limitación en Glue cuando no hay funciones definidas por el usuario (UDF).
- Esta versión corrige un problema que elimina los registros de contenedores mediante el servicio de agregación de registros de nodos antes de que el insertador de registros pudiera enviarlos a S3 en caso de retirada de YARN.
- Esta versión corrige un problema con las métricas del FairShare programador cuando Node Label está habilitado para Hadoop.
- Esta versión corrige un problema que afectaba al rendimiento de Spark cuando se establecía un valor `true` predeterminado para la configuración `spark.yarn.heterogeneousExecutors.enabled` en `spark-defaults.conf`.
- Esta versión corrige un problema que provocaba que la tarea de reducción no pudiera leer los datos de mezclas aleatorias. El problema provocaba errores en las consultas de Hive debido a un error de memoria dañada.
- Esta versión agrega un nuevo mecanismo de reintento al flujo de trabajo de escalado de clústeres para clústeres de EMR que ejecutan Presto o Trino. Esta mejora reduce el riesgo de que el cambio de tamaño del clúster se detenga indefinidamente debido a una única operación de cambio de

tamaño con errores. También mejora la utilización del clúster, ya que el clúster se escala y reduce verticalmente más rápido.

- Esta versión mejora la lógica de reducción vertical del clúster para que el clúster no intente reducir verticalmente los nodos principales por debajo de la configuración del factor de replicación de HDFS del clúster. Esto se ajusta a sus requisitos de redundancia de datos y reduce la posibilidad de que una operación de escalado se detenga.
- El daemon de administración de registros se ha actualizado para identificar todos los registros que están en uso activo con controladores de archivos abiertos en el almacenamiento de instancias local y los procesos asociados. Esta actualización garantiza que Amazon EMR elimine correctamente los archivos y recupere espacio de almacenamiento una vez archivados los registros en Amazon S3.
- Esta versión incluye una mejora del daemon de administración de registros que elimina los directorios escalonados vacíos y no utilizados del sistema de archivos del clúster local. Un número excesivamente elevado de directorios vacíos puede reducir el rendimiento de los daemons de Amazon EMR y provocar una sobreutilización del disco.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Malasia), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyy Derabad), EE. UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Malasia), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este), (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (AWS GovCloud Oregón), México (Centro), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Asia Pacífico (Taipéi), Canadá (Centro), Canadá Oeste (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512.0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Malasia), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Malasia) , Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Centro), Canadá Oeste (Calgary), (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China AWS GovCloud (Pekín), China (Ningxia), Israel (Tel Aviv), México (Centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich), Asia Pacífico (Malasia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 131.0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 124.0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia), Canadá Oeste (Calgary)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 218.0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 116.0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 101.0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 020.1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 012.1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda) , Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Hyderaba d), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Israel (Tel Aviv), AWS GovCloud (EEUU-

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Oeste), AWS GovCloud (EEUU-Este), China (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 906.0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

Versión 6.11.0

Las siguientes notas de la versión incluyen información sobre la versión 6.11.0 de Amazon EMR. Los cambios son respecto a la versión 6.10.0. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Nuevas características

- Amazon EMR 6.11.0 es compatible con Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0 y PrestoDB 0.279-amzn-0.

Cambios, mejoras y problemas resueltos

- Con Amazon EMR 6.11.0, el conector de DynamoDB se ha actualizado a la versión 5.0.0. La versión 5.0.0 usa. AWS SDK for Java 2.x En las versiones anteriores se utilizaba la versión AWS SDK para Java 1.x. Como resultado de esta actualización, se recomienda encarecidamente que pruebe el código antes de utilizar el conector de DynamoDB con Amazon EMR 6.11.
- Cuando el conector de DynamoDB para Amazon EMR 6.11.0 llama al servicio DynamoDB, utiliza el valor de región que usted proporciona para la propiedad `dynamodb.endpoint`. Se recomienda que también configure `dynamodb.region` cuando utilice `dynamodb.endpoint` y que ambas propiedades tengan la misma Región de AWS. Si lo usa `dynamodb.endpoint` y no lo configura`dynamodb.region`, el conector de DynamoDB para Amazon EMR 6.11.0 devolverá una excepción de región no válida e intentará conciliar la Región de AWS información del servicio de metadatos de instancias de Amazon (IMDS). EC2 Si el conector no puede recuperar la región del IMDS, el valor predeterminado será Este de EE. UU. (Norte de Virginia) (`us-east-1`). El siguiente error es un ejemplo de la excepción de región no válida que puede producirse si no configura la `dynamodb.region` propiedad correctamente: `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` para obtener más información sobre las clases afectadas por la AWS SDK para Java actualización a la versión 2.x, consulte la confirmación de actualización AWS SDK para Java de la versión 1.x a la 2.x (#175) en el GitHub repositorio del conector Amazon EMR - DynamoDB.
- Esta versión corrige un problema por el que los datos de las columnas se convierten en `NULL` cuando se utiliza Delta Lake para almacenar los datos de la tabla Delta en Amazon S3 tras una operación de cambio de nombre de columna. Para más información sobre esta característica experimental de Delta Lake, consulte Operación de cambio de nombre de columnas en la Guía del usuario de Delta Lake.
- La versión 6.11.0 corrige un problema que podía producirse al crear un nodo de periferia mediante la replicación de uno de los nodos principales de un clúster con varios nodos principales. El nodo de periferia replicado podría provocar retrasos en las operaciones de reducción vertical o provocar un uso elevado de la memoria en los nodos principales. Para obtener más información sobre cómo

crear un nodo perimetral para comunicarse con su clúster de EMR, consulte Edge Node Creator en el `aws-samples` repositorio de. GitHub

- La versión 6.11.0 mejora el proceso de automatización que Amazon EMR utiliza para volver a montar los volúmenes de Amazon EBS en una instancia tras un reinicio.
- La versión 6.11.0 corrige un problema que provocaba brechas intermitentes en las métricas de Hadoop que Amazon EMR publica en Amazon. CloudWatch
- La versión 6.11.0 corrige un problema con los clústeres de EMR por el que se interrumpe una actualización del archivo de configuración de YARN que contiene la lista de nodos excluidos del clúster debido a una sobreutilización del disco. La actualización incompleta dificulta las futuras operaciones de reducción vertical de los clústeres. Esta versión garantiza que el clúster se mantenga en buen estado y que las operaciones de escalado funcionen según lo previsto.
- El tamaño del volumen raíz predeterminado ha aumentado a 15 GB en la versión 6.10.0 y posteriores de Amazon EMR. Las versiones anteriores tienen un tamaño de volumen raíz predeterminado de 10 GB.
- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. Este enfoque también puede provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` se establece como `false` en `yarn-site.xml` para resolver este problema.

  Si bien la corrección soluciona los problemas que ha introducido YARN-9608, es posible que los trabajos de Hive presenten errores debido a la pérdida de datos de mezclas aleatorias en los clústeres que tienen habilitado el escalado administrado. Se ha mitigado ese riesgo en esta versión al configurar también `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` para las cargas de trabajo de Hive. Esta configuración solo está disponible con la versión 6.11.0 y posteriores de Amazon EMR.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

Versión 6.10.0

Las siguientes notas de la versión incluyen información sobre la versión 6.10.0 de Amazon EMR. Los cambios son respecto a la versión 6.9.0. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Nuevas características

- Amazon EMR 6.10.0 es compatible con Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403 y PrestoDB 0.278.1.
- Amazon EMR 6.10.0 incluye un conector Trino-Hudi nativo que proporciona acceso de lectura a los datos de las tablas de Hudi. Puede activar el conector con `trino-cli --catalog hudi` y configurarlo según sus requisitos con `trino-connector-hudi`. La integración nativa con Amazon EMR significa que ya no necesitará utilizar `trino-connector-hive` para consultar tablas de Hudi. Para obtener una lista de las configuraciones compatibles con el nuevo conector, consulte la página Hudi connector de la documentación de Trino.
- Las versiones 6.10.0 y posteriores de Amazon EMR admiten la integración de Apache Zeppelin con Apache Flink. Para obtener más información, consulte Trabajar con trabajos de Flink de Zeppelin en Amazon EMR.

Problemas conocidos

- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. Este enfoque también puede provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  Para solucionar este problema en Amazon EMR 6.10.0, puede establecer el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` en `yarn-site.xml`. En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, la configuración se establece en `false` de manera predeterminada para resolver este problema.

Cambios, mejoras y problemas resueltos

- Amazon EMR 6.10.0 elimina la dependencia en `minimal-json.jar` de la integración de Amazon Redshift para Apache Spark y agrega automáticamente los archivos JAR necesarios relacionados con Spark-Redshift a la ruta de clases de los ejecutores de Spark: `spark-redshift.jar`, `spark-avro.jar` y `RedshiftJDBC.jar`.
- La versión 6.10.0 mejora el daemon de administración de registros en el clúster para supervisar las carpetas de registro adicionales del clúster de EMR. Esta mejora minimiza los escenarios de sobreutilización del disco.
- La versión 6.10.0 reinicia automáticamente el daemon de administración de registros en el clúster cuando se detiene. Esta mejora reduce el riesgo de que los nodos parezcan estar en mal estado debido a la sobreutilización del disco.
- Amazon EMR 6.10.0 admite puntos de conexión regionales para la asignación de usuarios de EMRFS.
- El tamaño del volumen raíz predeterminado ha aumentado a 15 GB en la versión 6.10.0 y posteriores de Amazon EMR. Las versiones anteriores tienen un tamaño de volumen raíz predeterminado de 10 GB.
- La versión 6.10.0 corrige un problema que provocaba que los trabajos de Spark se detuvieran cuando todos los ejecutores de Spark restantes se encontraban en un host de retirada con el administrador de recursos de YARN.
- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Canadá Oeste (Calgary), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Zúrich), Europa (Milán), Europa (España), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pací fico (Yakarta), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU)

Versión 6.9.0

Las siguientes notas de la versión incluyen información sobre la versión 6.9.0 de Amazon EMR. Los cambios son respecto a la versión 6.8.0 de Amazon EMR. Para obtener información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Nuevas características

- La versión 6.9.0 de Amazon EMR es compatible con Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398 y Tez 0.10.2.
- La versión 6.9.0 de Amazon EMR incluye una nueva aplicación de código abierto, Delta Lake 2.1.0.
- La integración de Amazon Redshift para Apache Spark se incluye en las versiones 6.9.0 y posteriores de Amazon EMR. La integración nativa, que anteriormente era una herramienta de código abierto, es un conector de Spark que puede utilizar para crear aplicaciones de Apache Spark que leen y escriben datos en Amazon Redshift y Amazon Redshift sin servidor. Para obtener más información, consulte Uso de la integración de Amazon Redshift para Apache Spark con Amazon EMR .
- La versión 6.9.0 de Amazon EMR agrega compatibilidad para archivar registros en Amazon S3 durante la reducción vertical del clúster. Anteriormente, solo podía archivar archivos de registro en Amazon S3 durante la terminación del clúster. Esta nueva capacidad garantiza que los archivos de registro generados en el clúster persistan en Amazon S3 incluso después de terminar el nodo. Para más información, consulte Configurar el registro y la depuración de un clúster.
- Para admitir consultas de ejecución prolongada, Trino ahora incluye un mecanismo de ejecución tolerante a errores. La ejecución tolerante a errores mitiga los errores de las consultas al volver a intentar las consultas con errores o las tareas que las componen.
- Puede utilizar Apache Flink en Amazon EMR para el procesamiento unificado `BATCH` y `STREAM` de tablas de Apache Hive o metadatos de cualquier origen de tabla de Flink, como Iceberg, Kinesis o Kafka. Puedes especificar el catálogo de datos de AWS Glue como el metaalmacén de Flink mediante la API AWS Management Console AWS CLI, o Amazon EMR. Para obtener más información, consulte Configuración de Flink en Amazon EMR.
- Ahora puede especificar las funciones de tiempo de ejecución AWS Identity and Access Management (IAM) y el control de acceso AWS Lake Formation basado para las consultas de Apache Spark, Apache Hive y Presto en Amazon EMR en clústeres EC2 con Amazon AI Studio. SageMaker Para más información, consulte Configuración de roles en tiempo de ejecución para los pasos de Amazon EMR.

Problemas conocidos

- En el caso de la versión 6.9.0 de Amazon EMR, Trino no funciona en clústeres habilitados para Apache Ranger. Si tiene que utilizar Trino con Ranger, contacte con Soporte.
- Si utiliza la integración de Amazon Redshift para Apache Spark y tiene un valor de time, timetz, timestamp o timestamptz con una precisión de microsegundos en formato Parquet, el conector redondea los valores de tiempo al valor de milisegundos más cercano. Como solución alternativa, utilice el parámetro `unload_s3_format` de formato de descarga de texto.
- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Las conexiones a los clústeres de Amazon EMR desde Amazon SageMaker AI Studio pueden fallar de forma intermitente con un código de respuesta 403 Forbidden. Este error se produce cuando la configuración del rol de IAM en el clúster tarda más de 60 segundos. Como solución alternativa, puede instalar una revisión de Amazon EMR para permitir los reintentos y aumentar el tiempo de espera a un mínimo de 300 segundos. Siga los siguientes pasos para aplicar la acción de arranque al lanzar el clúster.

  1. Descargue el script de arranque y los archivos RPM del siguiente Amazon S3 URIs.

     ```
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
     ```

```
s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-
secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
```

2. Cargue los archivos del paso anterior a un bucket de Amazon S3 de su propiedad. El depósito debe estar en el mismo Región de AWS lugar en el que planea lanzar el clúster.
3. Incluya la siguiente acción de arranque cuando inicie el clúster de EMR. Sustituya *bootstrap_URI* y *RPM_URI* por la correspondiente URIs de Amazon S3.

```
--bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
```

- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-
actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-
actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- Apache Flink proporciona FileSystem conectores S3 FileSystem y Hadoop nativos, que permiten a las aplicaciones crear FileSink y escribir los datos en Amazon S3. Esto FileSink falla con una de las dos excepciones siguientes.

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
 supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError:
 org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/
```

```
apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/
RetryPolicy;
                                        at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-
yarn-common-3.3.3-amzn-0.jar:?]
```

Como solución alternativa, puede instalar una revisión de Amazon EMR, que soluciona el problema anterior en Flink. Siga los siguientes pasos para aplicar la acción de arranque al iniciar el clúster.

1. Descargue el archivo flink-rpm en su bucket de Amazon S3. La ruta de su RPM es `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`.
2. Descargue el script de arranque y los archivos RPM de Amazon S3 con los siguientes URI. *regionName*Sustitúyalo por el Región de AWS lugar en el que planea lanzar el clúster.

   ```
   s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-
   rpms.sh
   ```

3. Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. En Amazon EMR 6.8.0 y 6.9.0, este enfoque también podría provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

   Con Amazon EMR 6.10.0, hay una solución alternativa para este problema: puede establecer el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` en `yarn-site.xml`. En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, la configuración se establece en `false` de manera predeterminada para resolver este problema.

Cambios, mejoras y problemas resueltos

- En el caso de la versión 6.9.0 y posteriores de Amazon EMR, todos los componentes instalados por Amazon EMR que utilizan bibliotecas Log4j utilizan la versión 2.17.1 o posteriores de Log4j.
- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. La versión 6.9.0 de Amazon EMR corrige este problema.

- Amazon EMR 6.9.0 agrega compatibilidad limitada para el control de acceso basado en Lake Formation con Apache Hudi al leer datos con Spark SQL. La compatibilidad es para consultas SELECT que utilizan Spark SQL y se limita al control de acceso en el nivel de columnas. Para más información, consulte Hudi y Lake Formation.
- Cuando utiliza Amazon EMR 6.9.0 para crear un clúster de Hadoop con etiquetas de nodo habilitadas, la API de métricas de YARN devuelve información agregada en todas las particiones, en lugar de la partición predeterminada. Para más información, consulte YARN-11414.
- Con la versión 6.9.0 de Amazon EMR, hemos actualizado Trino a la versión 398, que utiliza Java 17. La versión anterior compatible de Trino para Amazon EMR 6.8.0 era Trino 388 y se ejecutaba en Java 11. Para más información sobre este cambio, consulte las actualizaciones de Trino a Java 17 en el blog de Trino.
- Esta versión corrige un problema de desajuste de la secuencia de tiempo entre Apache BigTop y Amazon EMR EC2 en la secuencia de inicio del clúster. Este desajuste de la secuencia de tiempo se produce cuando un sistema intenta llevar a cabo dos o más operaciones al mismo tiempo en lugar de hacerlas en la secuencia correcta. Como resultado, algunas configuraciones de clústeres experimentaron tiempos de espera de inicio de instancias y tiempos de inicio de clústeres más lentos.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Canadá (Centro), Canadá Occidental (Calgary), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), Canadá Oeste (Calgary) AWS

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud , (EEUU-Este) AWS GovCloud , (EEUU-Oeste), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,14,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,14,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Europa (Zúrich)
2.0.2024 001.0	4,14352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv), Canadá Oeste (Calgary)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro), Israel (Tel Aviv)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	12 de enero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

Versión 6.8.0

Las siguientes notas de la versión incluyen información sobre la versión 6.8.0 de Amazon EMR. Los cambios son respecto a la versión 6.7.0.

Nuevas características

- La función de pasos de Amazon EMR ahora es compatible con los clientes y puntos de conexión de Apache Livy. JDBC/ODBC Para más información, consulte Configuración de roles en tiempo de ejecución para los pasos de Amazon EMR.
- La versión 6.8.0 de Amazon EMR viene con la versión 2.4.12 de Apache HBase . Con esta HBase versión, puede archivar y eliminar sus tablas. HBase El proceso de archivado de Amazon S3 cambia el nombre de todos los archivos de tabla al directorio de archivado. Esto puede ser un proceso largo y costoso. Ahora puede omitir el proceso de archivado y rechazar y eliminar rápidamente tablas grandes. Para obtener más información, consulte Uso del HBase shell.

Problemas conocidos

- Hadoop 3.3.3 ha introducido un cambio en YARN (YARN-9608) que mantiene los nodos donde se ejecutaron los contenedores en un estado de retirada hasta que se complete la aplicación. Este cambio garantiza que los datos locales, como los datos de mezclas aleatorias, no se pierdan y que no sea necesario volver a ejecutar el trabajo. En Amazon EMR 6.8.0 y 6.9.0, este enfoque también podría provocar una infrautilización de los recursos en los clústeres con o sin el escalado administrado habilitado.

  Con Amazon EMR 6.10.0, hay una solución alternativa para este problema: puede establecer el valor de `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` como `false` en `yarn-site.xml`. En las versiones 6.11.0 y posteriores de Amazon EMR, así como en las 6.8.1, 6.9.1 y 6.10.1, la configuración se establece en `false` de manera predeterminada para resolver este problema.

Cambios, mejoras y problemas resueltos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leían las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, Amazon EMR producía un `NoSuchMethodError`. La versión 6.8.0 de Amazon EMR corrige este problema.
- Amazon EMR 6.8.0 viene con Apache Hudi 0.11.1; sin embargo, los clústeres de Amazon EMR 6.8.0 también son compatibles con el código abierto `hudi-spark3.3-bundle_2.12` de Hudi 0.12.0.
- La versión 6.8.0 de Amazon EMR incluye la versión 3.3.0 de Apache Spark. Esta versión de Spark utiliza Apache Log4j 2 y el archivo `log4j2.properties` para configurar Log4j en los procesos

de Spark. Si utiliza Spark en el clúster o crea clústeres de EMR con parámetros de configuración personalizados y desea actualizar a la versión 6.8.0 de Amazon EMR, debe migrar al nuevo formato de clave y clasificación de configuración `spark-log4j2` de Apache Log4j 2. Para obtener más información, consulte Migración de Apache Log4j 1.x a Log4j 2.x.

- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

> **Note**
>
> Esta versión ya no recibe actualizaciones automáticas de la AMI, ya que ha sido sustituida por 1 versión de revisión adicional. La versión de revisión se indica con el número que sigue al segundo punto decimal (`6.8.1`). Para comprobar si está utilizando la versión de revisión más reciente, consulte las versiones disponibles en la Guía de versiones o consulte el menú desplegable Versiones de Amazon EMR cuando cree un clúster en la consola o utilice la acción de la API `ListReleaseLabels` o la CLI `list-release-labels`. Para recibir actualizaciones sobre las nuevas versiones, suscríbase a la fuente RSS de la página Novedades.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			(Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Asia Pacífico (Melbourne), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este) AWS GovCloud , (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Israel (Tel Aviv), Canadá (Centro), Europa (España), China (Ningxia), Europa (Zúrich)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Melbourne), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), AWS GovCloud (EEUU-Este), AWS

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			GovCloud (EEUU-Oes te), China (Pekín), China (Ningxia), Israel (Tel Aviv)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Bombay), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos), Asia Pacífico (Melbourne),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Europa (Zúrich)
2.0.2024 001.0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Árabes Unidos), Europa (España), Europa (Zúrich), Asia Pacífico (Melbourne), Israel (Tel Aviv)
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pací fico (Osaka), Asia-Pací fico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), Asia-Pacífico (Melbourne), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 119,1	4,14.301	3 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	22 de diciembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 912,1	4,14,291	6 de septiembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

Problemas conocidos

- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. Esto se debe a que Spark 3.2.0 establece

`spark.hadoopRDD.ignoreEmptySplits` en `true` de manera predeterminada. Como solución alternativa, establezca `spark.hadoopRDD.ignoreEmptySplits` explícitamente en `false`. La versión 6.9.0 de Amazon EMR corrige este problema.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Para obtener más información sobre el cronograma de lanzamiento, consulte el registro de cambios.

Versión 6.7.0

Las siguientes notas de la versión incluyen información sobre la versión 6.7.0 de Amazon EMR. Los cambios son respecto a la versión 6.6.0.

Fecha de lanzamiento inicial: 15 de julio de 2022

Nuevas características

- Amazon EMR ahora es compatible con Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272 y Trino 0.378.
- Admite controles de acceso basados en IAM Role y Lake Formation con pasos de EMR (Spark, Hive) para Amazon EMR en clústeres. EC2
- Admite las instrucciones de definición de datos de Apache Spark en clústeres habilitados para Apache Ranger. Esto ahora incluye compatibilidad para que las aplicaciones de Trino lean y escriban metadatos de Apache Hive en clústeres habilitados para Apache Ranger. Para más información, consulte Habilitar la gobernanza federada mediante Trino y Apache Ranger en Amazon EMR.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin), China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Oriente Medio (Emiratos Árabes Unidos), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), EE.UU. Este (N. Virginia), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Europa (Milán), Asia Pacífico (Tokio), (EE. UU. Este), (EE. UU. Oeste) AWS GovCloud , EE. UU. Oeste (Oregón) AWS GovCloud , EE. UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia Pacífico (Sídney)), Israel (Tel

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Aviv), Canadá (Centro), Europa (España), China (Ningxia), Europa (Zúrich)
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (España), Europa (Estocolmo), Europa (Zúrich), Asia Pacífico (Hong Kong), Asia Pacífico (Hyderabad), Asia Pacífico (Yakarta), Asia Pacífico (Bombay), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur), Asia Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Oriente Medio (Emiratos Árabes Unidos), Canadá (Central), (EE. UU. AWS GovCloud Este), AWS GovCloud (EEUU-Oes

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			te), China (Pekín), China (Ningxia), Israel (Tel Aviv)
2.0.2025 414.0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Medio Este (Emiratos Árabes Unidos),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Israel (Tel Aviv), Europa (Zúrich)
2.0.2024 001.0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia), Asia Pacífico (Hyderabad), Oriente Medio (Emiratos Árabes

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			Unidos), Europa (España), Europa (Zúrich)
2.0.2024 223.0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 131,0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 124,0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 218,0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 206,0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 116,0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 101,0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 020,1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 012,1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 906,0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4,114,314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4,14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 119,1	4,14.301	3 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	22 de diciembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 912,1	4,14,291	7 de octubre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2022 719,0	4,14,287	10 de agosto de 2022	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 606.1	4,14,281	15 de julio de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

Problemas conocidos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leen las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, se produce un error `NoSuchMethodError`

porque Amazon EMR utiliza un valor de `Hbase.compat.version` incorrecto. La versión 6.8.0 de Amazon EMR corrige este problema.

- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. Esto se debe a que Spark 3.2.0 establece `spark.hadoopRDD.ignoreEmptySplits` en `true` de manera predeterminada. Como solución alternativa, establezca `spark.hadoopRDD.ignoreEmptySplits` explícitamente en `false`. La versión 6.9.0 de Amazon EMR corrige este problema.
- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- La API `GetClusterSessionCredentials` no es compatible con los clústeres que se ejecutan en Amazon EMR 6.7 o versiones anteriores.

Versión 6.6.0

Las siguientes notas de la versión incluyen información sobre la versión 6.6.0 de Amazon EMR. Los cambios son respecto a la versión 6.5.0.

Fecha de lanzamiento inicial: 9 de mayo de 2022

Última actualización de la documentación: 15 de junio de 2022

Nuevas características

- Amazon EMR 6.6 ahora admite Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367 y PrestoDB 0.267.
- Al lanzar un clúster con la versión de parche más reciente de Amazon EMR 5.36, 6.6 o 7.0 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux 2023 o Amazon Linux 2 para la AMI de Amazon EMR predeterminada. Para más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR.

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2025 527.1	4.14.355-277.647.amzn2	19 de junio de 2025	Europa (Estocolmo), Oriente Medio (Baréin),

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
			China (Pekín), Asia Pacífico (Bombay), Europa (París), Asia Pacífico (Yakarta), EE.UU. Este (Ohio), África (Ciudad del Cabo), Europa (Irlanda), Europa (Fráncfort), Sudamérica (São Paulo), Asia Pacífico (Hong Kong), EE.UU. Este (Norte de Virginia) , Asia Pacífico (Seúl), Asia Pacífico (Osaka), Europa (Londres), Europa (Milán), Asia Pacífico (Tokio), (EEUU-Este) AWS GovCloud , (EEUU-Oeste) AWS GovCloud , EE.UU. Oeste (Oregón), EE.UU. Oeste (Norte de California), Asia Pacífico (Singapur), Asia-Pacífico (Sídney), Canadá (Centro), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (Milán), Europa (París), Europa (Estocolmo), Asia Pacífico (Hong Kong), Asia Pacífico (Yakarta), Asia Pacífico (Mumbai), Asia Pacífico (Osaka), Asia Pacífico (Seúl), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Tokio), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Central) , (EEUU-Este), (EEUU-Oeste) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 414,0	4.14.355-276.618.amzn2	12 de mayo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 321,0	4,14,355	09 de abril de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 305,0	4,114,355	18 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 220,0	4,14,355	8 de marzo de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 201,0	4,14,355	28 de febrero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 123,4	4,14,355	27 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2025 116.0	4,114,355	23 de enero de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 217,0	4,114,355	8 de enero de 2025	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 001,0	4,114,352	4 de octubre de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 816,0	4,14,350	21 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 809,0	4,14,349	20 de agosto de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 719,0	4,14,348	25 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 709,1	4.14.348	23 de julio de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 223,0	4,14,336	8 de marzo de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 131,0	4,14,336	14 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 124,0	4,14.336	7 de febrero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2024 109,0	4,1334	24 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 218,0	4,14,330	2 de enero de 2024	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 206,0	4,14.330	22 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 116,0	4,14328	11 de diciembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 101,0	4.14.327	16 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 020,1	4,14326	7 de noviembre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 012,1	4,14326	26 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 926,0	4,14,322	19 de octubre de 2023	EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio), EE.UU. Oeste (Norte de California), EE.UU. Oeste (Oregón), Europa (Estocolm o), Europa (Milán), Europa (Fráncfort), Europa (Irlanda), Europa (Londres), Europa (París), Asia Pacífico (Hong Kong), Asia Pacífico (Mumbai), Asia Pacífico (Tokio), Asia Pacífico (Seúl), Asia Pacífico (Osaka), Asia Pacífico (Singapur) Singapur), Asia-Pacífico (Sídney), Asia Pacífico (Yakarta), África (Ciudad del Cabo), Sudamérica (São Paulo), Oriente Medio (Baréin), Canadá (Centro), (EEUU-Oeste), (EEUU-Este) AWS GovCloud , China AWS GovCloud (Pekín), China (Ningxia)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2,0,2023 906,0	4,14,322	4 de octubre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 822,0	4,14,322	30 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 808.0	4,14320	24 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 727,0	4,14320	14 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (Fráncfor t), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 719,0	4,14320	2 de agosto de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Europa (Estocolmo), Europa (Milán), Europa (España), Europa (Fráncfort), Europa (Zúrich), Europa (Irlanda), Europa (Londres), Europa (París), Asia-Pacífico (Hong Kong), Asia-Pací fico (Bombay), Asia-Pací fico (Hyderabad), Asia-Pacífico (Tokio), Asia-Pací fico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), Asia-Pacífico (Yakarta), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin), Medio Oriente (EAU), Canadá (centro)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 628,0	4,14318	12 de julio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 612,0	4.14.314	23 de junio de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 504.1	4.14.313	16 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 418,0	4,14,311	3 de mayo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 404.1	4,14,311	18 de abril de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2023 404.0	4,14,311	10 de abril de 2023	Este de EE. UU. (Norte de Virginia), Europa (París)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 320,0	4,14309	30 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 307.0	4,14.305	15 de marzo de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 207,0	4,14,304	22 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2023 119,1	4,14.301	3 de febrero de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 210.1	4.14.301	22 de diciembre de 2023	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 912,1	4,14,291	7 de octubre de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)
2.0.2022 805.0	4,14,287	30 de agosto de 2022	`us-west-1`

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 719,0	4,14,287	10 de agosto de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 426,0	4,14,281	10 de junio de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

OsRelea Label (Versión de Amazon Linux)	Versión de kernel de Amazon Linux	Fecha de disponibilidad	Regiones compatibles
2.0.2022 406.1	4,14,275	2 de mayo de 2022	Este de EE. UU. (Norte de Virginia), Este de EE. UU. (Ohio), Oeste de EE. UU. (Norte de California), Oeste de EE. UU. (Oregón), Canadá (centro), Europa (Estocolmo), Europa (Irlanda), Europa (Londres), Europa (París), Europa (Fráncfort), Europa (Milán), Asia-Pacífico (Hong Kong), Asia-Pacífico (Bombay), Asia-Pacífico (Yakarta), Asia-Pacífico (Tokio), Asia-Pacífico (Seúl), Asia-Pacífico (Osaka), Asia-Pacífico (Singapur), Asia-Pacífico (Sídney), África (Ciudad del Cabo), América del Sur (São Paulo), Medio Oriente (Baréin)

- Con Amazon EMR 6.6 y versiones posteriores, las aplicaciones que utilizan Log4j 1.x y Log4j 2.x se actualizan para utilizar Log4j 1.2.17 (o posterior) y Log4j 2.17.1 (o posterior), respectivamente, y no requieren el uso de las acciones de arranque proporcionadas para mitigar los problemas de CVE.
- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y

posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.

- A partir de Amazon EMR 5.32.0 y 6.5.0, el tamaño dinámico del ejecutor para Apache Spark se habilita de forma predeterminada. Para activar o desactivar esta característica, puede utilizar el parámetro de configuración `spark.yarn.heterogeneousExecutors.enabled`.

Cambios, mejoras y problemas resueltos

- Amazon EMR reduce el tiempo de inicio del clúster hasta 80 segundos de promedio para los clústeres que utilizan la opción de AMI predeterminada de EMR y que solo instalan aplicaciones comunes, como Apache Hadoop, Apache Spark y Apache Hive.

Problemas conocidos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leen las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, se produce un error `NoSuchMethodError` porque Amazon EMR utiliza un valor de `Hbase.compat.version` incorrecto. La versión 6.8.0 de Amazon EMR corrige este problema.
- Cuando utiliza el conector de DynamoDB con Spark en las versiones 6.6.0, 6.7.0 y 6.8.0 de Amazon EMR, todas las lecturas de la tabla devuelven un resultado vacío, aunque la división de entrada haga referencia a datos que no están vacíos. Esto se debe a que Spark 3.2.0 establece `spark.hadoopRDD.ignoreEmptySplits` en `true` de manera predeterminada. Como solución alternativa, establezca `spark.hadoopRDD.ignoreEmptySplits` explícitamente en `false`. La versión 6.9.0 de Amazon EMR corrige este problema.
- En los clústeres de ejecución prolongada de Trino, Amazon EMR 6.6.0 habilita los parámetros de registro de recopilación de elementos no utilizados en el archivo jvm.config de Trino para obtener mejor información de los registros de recopilación de elementos no utilizados. Este cambio añade muchos registros de recolección de basura al archivo launcher.log (/var/log/trino/launcher.log). Si ejecuta clústeres de Trino en Amazon EMR 6.6.0, es posible que los nodos se queden sin espacio en disco después de que el clúster haya estado ejecutándose durante un par de días debido a los registros agregados.

La solución alternativa para este problema consiste en ejecutar el siguiente script como una acción de arranque para deshabilitar los parámetros de registro de recopilación de elementos no utilizados en jvm.config al crear o clonar el clúster para Amazon EMR 6.6.0.

```
#!/bin/bash
  set -ex
  PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
  sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.
- Con las versiones 5.36.0 y 6.6.0 a 6.9.0 de Amazon EMR, los componentes de servicio `SecretAgent` y `RecordServer` pueden sufrir una pérdida de datos de registro debido a una configuración incorrecta del patrón de nombres de archivo en las propiedades de Log4j2. Una configuración incorrecta hace que los componentes generen solo un archivo de registro por día. Cuando se aplica la estrategia de rotación, se sobrescribe el archivo existente en lugar de generar un nuevo archivo de registro como se esperaba. Como solución alternativa, utilice una acción de arranque para generar archivos de registro cada hora y agregue un entero de incremento automático en el nombre del archivo para gestionar la rotación.

  En el caso de las versiones 6.6.0 a 6.9.0 de Amazon EMR, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

En el caso de Amazon EMR 5.36.0, utilice la siguiente acción de arranque al lanzar un clúster.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

Versión 5.35.0

Esta es la nota de la versión 5.35.0 de Amazon EMR.

Las siguientes notas de la versión incluyen información sobre la versión 5.35.0 de Amazon EMR. Los cambios son respecto a la versión 5.34.0.

Fecha de lanzamiento inicial: 30 de marzo de 2022

Nuevas características

- Las aplicaciones de Amazon EMR, versión 5.35, que utilizan Log4j 1.x y Log4j 2.x se actualizan para utilizar Log4j 1.2.17 (o posterior) y Log4j 2.17.1 (o posterior), respectivamente, y no requieren el uso de acciones de arranque para mitigar los problemas de CVE de las versiones anteriores. Consulte Método para mitigar CVE-2021-44228.

Cambios, mejoras y problemas resueltos

Cambios en Flink

Cambio de tipo	Descripción
Actualizaciones	• Actualice la versión de Flink a la 1.14.2. • log4j se actualizó a la versión 2.17.1.

Cambios en Hadoop

Cambio de tipo	Descripción
Portabilidad con versiones anteriores de código abierto de Hadoop desde EMR 5.34.0	• YARN-10438: Maneja un ContainerID nulo en el cliente # () RMService getContainerReport • YARN-7266: los subprocesos del controlador de eventos del servidor Timeline están bloqueados • YARN-10438: ATS 1.5 no se inicia si los archivos están dañados o faltan RollingLevelDb • HADOOP-13500: sincronización de iteración del objeto de propiedades de configuración • YARN-10651: se estrelló con el NPE activado. CapacityScheduler AbstractYarnScheduler updateNodeResource() • HDFS-12221: Sustituya los xerces en XmlEditsVisitor • HDFS-16410: análisis XML inseguro en OfflineEditsXmlLoader
Cambios y correcciones de Hadoop	• Tomcat utilizado en KMS y HTTPFS se ha actualizado a la versión 8.5.75 • En la FileSystemOptimizedCommitter versión 2, el marcador de éxito se escribía en la ruta de salida de CommitJob definida al crear el confirmador. Como las rutas de salida de commitJob y en el nivel de tarea pueden diferir, se corrigió la ruta para usar la definida en los archivos de manifiesto. En el caso de los trabajos de Hive, esto hace que el marcador de éxito se escriba correctamente al realizar operaciones como la partición dinámica o UNION ALL.

Cambios en Hive

Cambio de tipo	Descripción
Hive se ha actualizado a la versión 2.3.9 de código abierto, que incluye estas correcciones de JIRA	• HIVE-17155: findConfFile () en HiveConf .java tiene algunos problemas con la ruta de configuración • HIVE-24797: deshabilite la validación de valores predeterminados al analizar los esquemas de Avro • HIVE-21563: Mejore el rendimiento de la tabla # deshabilitándola una vez getEmptyT able registerAllFunctions • HIVE-18147: Las pruebas pueden fallar con java.net. BindException: La dirección ya está en uso • HIVE-24608: cambie de nuevo a get_table en el cliente de HMS para Hive 2.3.x • HIVE-21200: Vectorización: columna de fecha que arroja java.lang. Unsupport edOperationException para parquet • HIVE-19228: elimine el uso de commons-h ttpclient 3.x
Backports de código abierto de Hive desde EMR 5.34.0	• HIVE-19990: la consulta con un literal de intervalo en la condición de unión falla • HIVE-25824: actualice branch-2.3 a log4j 2.17.0 • TEZ-4062: la programación de intentos especulativos debe anularse cuando la tarea se haya completado • TEZ-4108: en condiciones de carrera NullPointerException durante una ejecución especulativa

Cambio de tipo	Descripción
	• TEZ-3918: la configuración de tez.task.log.level no funciona
Actualizaciones y correcciones de Hive	• Actualice la versión de Log4j a la 2.17.1 • Actualice la versión de ORC a la 1.4.3 • Se corrigió un punto muerto debido a un hilo de penalización en ShuffleScheduler
Nuevas características	• Se agregó una característica para imprimir consultas de Hive en los registros de AM. Esta opción está deshabilitada de forma predeterminada. Marca/Conf.: `tez.am.emr.print.hive.query.in.log` . Estado (predeterminado): FALSO.

Cambios en Oozie

Cambio de tipo	Descripción
Backports de código abierto de Oozie desde EMR 5.34.0	• OOZIE-3652: El lanzador de Oozie debería volver a intentar enumerar los directorios cuando aparezca NoSuchFileException

Cambios en Pig

Cambio de tipo	Descripción
Actualizaciones	• log4j se actualizó a la versión 1.2.17.

Problemas conocidos

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:

- Se analizan dos o más particiones de la misma tabla.
- Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
- El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 5.34.0

Las siguientes notas de la versión incluyen información sobre la versión 5.34.0 de Amazon EMR. Los cambios son respecto a la versión 5.33.1.

Fecha de lanzamiento inicial: 20 de enero de 2022

Fecha de lanzamiento: 21 de marzo de 2022

Nuevas características

- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.
- [Hudi] Mejoras para simplificar la configuración de Hudi. El control de simultaneidad optimista se deshabilita de forma predeterminada.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Anteriormente, el reinicio manual del administrador de recursos en un clúster multimaestro provocaba que los daemons en el clúster de Amazon EMR, como ZooKeeper, recargaran todos los nodos perdidos o retirados previamente en el archivo znode de ZooKeeper. Esto provocaba que se superaran los límites predeterminados en ciertas situaciones. Amazon EMR ahora elimina del archivo de ZooKeeper los registros de nodos perdidos o retirados que tengan más de una hora de antigüedad y se han incrementado los límites internos.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Se ha actualizado Zeppelin a la versión 0.10.0.
- Corrección de Livy: actualizado a la versión 0.7.1

- Mejora del rendimiento de Spark: los ejecutores heterogéneos se deshabilitan cuando se anulan ciertos valores de configuración de Spark en EMR 5.34.0.
- Los servidores WebHDFS y HttpFS están deshabilitados de forma predeterminada. Puede volver a habilitar WebHDFS mediante la configuración de Hadoop, `dfs.webhdfs.enabled`. El servidor HttpFS se puede iniciar mediante `sudo systemctl start hadoop-httpfs`.

Problemas conocidos

- La característica Cuadernos de Amazon EMR que se utiliza con la suplantación de usuarios de Livy no funciona porque HttpFS está deshabilitado de forma predeterminada. En este caso, el cuaderno de EMR no puede conectarse al clúster que tiene habilitada la suplantación de Livy. La solución alternativa consiste en iniciar el servidor HttpFS antes de conectar el cuaderno de EMR al clúster mediante `sudo systemctl start hadoop-httpfs`.
- Las consultas de Hue no funcionan en Amazon EMR 6.4.0 porque el servidor HttpFS de Apache Hadoop está deshabilitado de forma predeterminada. Para usar Hue en Amazon EMR 6.4.0, inicie manualmente el servidor HttpFS en el nodo principal de Amazon EMR mediante `sudo systemctl start hadoop-httpfs` o utilice un paso de Amazon EMR.
- La característica Cuadernos de Amazon EMR que se utiliza con la suplantación de usuarios de Livy no funciona porque HttpFS está deshabilitado de forma predeterminada. En este caso, el cuaderno de EMR no puede conectarse al clúster que tiene habilitada la suplantación de Livy. La solución alternativa consiste en iniciar el servidor HttpFS antes de conectar el cuaderno de EMR al clúster mediante `sudo systemctl start hadoop-httpfs`.
- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 6.5.0

Las siguientes notas de la versión incluyen información sobre la versión 6.5.0 de Amazon EMR. Los cambios son respecto a la versión 6.4.0.

Fecha de lanzamiento inicial: 20 de enero de 2022

Fecha de lanzamiento: 21 de marzo de 2022

Nuevas características

- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.
- A partir de Amazon EMR 5.32.0 y 6.5.0, el tamaño dinámico del ejecutor para Apache Spark se habilita de forma predeterminada. Para activar o desactivar esta característica, puede utilizar el parámetro de configuración `spark.yarn.heterogeneousExecutors.enabled`.
- Compatibilidad con el formato de tabla abierto Apache Iceberg para conjuntos de datos analíticos muy grandes.
- Support para ranger-trino-plugin 2.0.1-amzn-1
- Compatibilidad con toree 0.5.0

Cambios, mejoras y problemas resueltos

- La versión 6.5 de Amazon EMR ahora es compatible con Apache Iceberg 0.12.0 y ofrece mejoras en el tiempo de ejecución con Tiempo de ejecución de Amazon EMR para Apache Spark, Tiempo de ejecución de Amazon EMR para Presto y Tiempo de ejecución de Amazon EMR para Apache Hive.

- Apache Iceberg es un formato de tabla abierto para grandes conjuntos de datos en Amazon S3 y proporciona un rendimiento de consultas rápido en tablas grandes, confirmaciones atómicas, escrituras simultáneas y evolución de tablas compatible con SQL. A partir de Amazon EMR 6.5, puede usar Apache Spark 3.1.2 con el formato de tabla de Iceberg.
- Apache Hudi 0.9 agrega compatibilidad con DDL y DML de Spark SQL. Esto le permite crear y upsert tablas de Hudi utilizando únicamente instrucciones SQL. Apache Hudi 0.9 también incluye mejoras en el rendimiento de las consultas y del escritor.
- El Tiempo de ejecución de Amazon EMR para Apache Hive mejora el rendimiento de Apache Hive en Amazon S3 al eliminar las operaciones de cambio de nombre durante las operaciones de preparación y mejora el rendimiento de los comandos de comprobación de metaalmacén (MSCK) que se utilizan para reparar tablas.

Problemas conocidos

- Cuando las versiones 6.5.0, 6.6.0 o 6.7.0 de Amazon EMR leen las tablas de Apache Phoenix a través del intérprete de comandos de Apache Spark, se produce un error `NoSuchMethodError` porque Amazon EMR utiliza un valor de `Hbase.compat.version` incorrecto. La versión 6.8.0 de Amazon EMR corrige este problema.
- Los clústeres de paquetes de HBase en alta disponibilidad (HA) no se aprovisionan con el tamaño de volumen y el tipo de instancia predeterminados. La solución alternativa para este problema consiste en aumentar el tamaño del volumen raíz.
- Para utilizar las acciones de Spark con Apache Oozie, debe agregar la siguiente configuración al archivo `workflow.xml` de Oozie. De lo contrario, varias bibliotecas críticas, como Hadoop y EMRFS, no aparecerán en la ruta de clases de los ejecutores de Spark que lance Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.

- El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter / (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 6.4.0

Las siguientes notas de la versión incluyen información sobre la versión 6.4.0 de Amazon EMR. Los cambios son respecto a la versión 6.3.0.

Fecha de lanzamiento inicial: 20 de septiembre de 2021

Fecha de lanzamiento: 21 de marzo de 2022

Aplicaciones compatibles

- AWS SDK para Java versión 1.12.31
- CloudWatch Sink versión 2.2.0
- Conector de DynamoDB, versión 4.16.0
- EMRFS, versión 2.47.0
- Amazon EMR Goodies, versión 3.2.0
- Conector de Kinesis para Amazon EMR, versión 3.5.0
- Servidor de registros de Amazon EMR, versión 2.1.0
- Scripts de Amazon EMR, versión 2.5.0
- Flink, versión 1.13.1
- Versión de Ganglia 3.7.2
- AWS Cliente Glue Hive Metastore versión 3.3.0
- Hadoop, versión 3.2.1-amzn-4
- HBase versión 2.4.4-amzn-0
- HBase-operator-tools 1.1.0

- HCatalog versión 3.1.2-amzn-5
- Hive, versión 3.1.2-amzn-5
- Hudi versión 0.8.0-amzn-0
- Hue versión 4.9.0
- Java JDK, versión Corretto-8.302.08.1 (compilación 1.8.0_302-b08)
- JupyterHub versión 1.4.1
- Livy versión 0.7.1-incubating
- MXNet versión 1.8.0
- Oozie versión 5.2.1
- Phoenix versión 5.1.2
- Pig, versión 0.17.0
- Presto, versión 0.254.1-amzn-0
- Trino, versión 359
- Apache Ranger KMS (cifrado transparente multimaestro), versión 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker Spark SDK versión 1.4.1
- Scala, versión 2.12.10 (máquina virtual de servidor OpenJDK de 64 bits, Java 1.8.0_282)
- Spark, versión 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop, versión 1.4.7
- TensorFlow versión 2.4.1
- Tez, versión 0.9.2
- Zeppelin versión 0.9.0
- ZooKeeper versión 3.5.7
- Conectores y controladores: conector de DynamoDB 4.16.0

Nuevas características

- [Escalado administrado] Optimización del escalado administrado de datos de mezclas aleatorias de Spark: para las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y

posteriores de EMR, el escalado administrado ahora es compatible con los datos de mezclas aleatorias de Spark (datos que Spark redistribuye entre las particiones para realizar operaciones específicas). Para más información sobre las operaciones de mezclas aleatorias, consulte Uso del escalado administrado de EMR en Amazon EMR en la Guía de administración de Amazon EMR y Spark Programming Guide.

- En los clústeres de Amazon EMR habilitados para Apache Ranger, puede utilizar Apache Spark SQL para insertar datos o actualizar las tablas del metaalmacén de Apache Hive mediante `INSERT INTO`, `INSERT OVERWRITE` y `ALTER TABLE`. Al utilizar ALTER TABLE con Spark SQL, la ubicación de una partición debe ser el directorio secundario de la ubicación de una tabla. Amazon EMR no admite actualmente la inserción de datos en una partición en la que la ubicación de la partición es diferente de la ubicación de la tabla.
- PrestoSQL se ha renombrado como Trino.
- Hive: la ejecución de consultas SELECT sencillas con la cláusula LIMIT se acelera al detener la ejecución de las consultas en cuanto se obtiene el número de registros mencionado en la cláusula LIMIT. Las consultas SELECT simples son consultas que no tienen GROUP BY u ORDER BY por cláusula o consultas que no tienen una etapa reductora. Por ejemplo, `SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`.

Control de simultaneidad de Hudi

- Hudi ahora es compatible con el control de simultaneidad optimista (OCC), que se puede utilizar con operaciones de escritura como UPSERT e INSERT para permitir cambios desde varios escritores a la misma tabla de Hudi. Se trata de un OCC de nivel de archivo, por lo que dos confirmaciones (o escritores) pueden escribir en la misma tabla, siempre que sus cambios no entren en conflicto. Para más información, consulte Control de simultaneidad de Hudi.
- Los clústeres de Amazon EMR tienen instalado ZooKeeper, que se puede utilizar como proveedor de bloqueos para OCC. Para facilitar el uso de esta característica, los clústeres de Amazon EMR tienen las siguientes propiedades preconfiguradas:

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

Para habilitar el OCC, debe configurar las siguientes propiedades con sus opciones de trabajo de Hudi o en el nivel de clúster mediante la API de configuraciones de Amazon EMR:

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Hudi Monitoring: CloudWatch integración de Amazon para informar sobre las métricas de Hudi

- Amazon EMR admite la publicación de Hudi Metrics en Amazon. CloudWatch Se habilita mediante el establecimiento de las siguientes configuraciones obligatorias:

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- Las siguientes son configuraciones opcionales de Hudi que se pueden cambiar:

Opción	Descripción	Valor
hoodie.metrics.cloudwatch.report.period.seconds	Frecuencia (en segundos) con la que se reportan las métricas a Amazon CloudWatch	El valor predeterminado es 60 segundos, lo cual está bien para la resolución predeterminada de un minuto que ofrece Amazon CloudWatch
hoodie.metrics.cloudwatch.metric.prefix	Prefijo que se agregará al nombre de cada métrica	El valor predeterminado es vacío (sin prefijo)
hoodie.metrics.cloudwatch.namespace	Espacio de CloudWatch nombres de Amazon en el que se publican las métricas	El valor predeterminado es Hudi
hoodie.metrics.cloudwatch.maxDatumsPerSolicitud	Número máximo de datos que se pueden incluir en una solicitud a Amazon CloudWatch	El valor predeterminado es 20, que es el mismo que el CloudWatch valor predeterminado de Amazon

Soporte y mejoras en las configuraciones de Amazon EMR Hudi

- Los clientes ahora pueden aprovechar la API de configuraciones de EMR y la característica de reconfiguración para configurar los ajustes de Hudi en el nivel de clúster. Se ha introducido un nuevo soporte de configuración basado en archivos a través de/etc/hudi/conf/hudi-defaults.conf, similar al de otras aplicaciones como Spark, Hive, etc. EMR configura algunos valores predeterminados para mejorar la experiencia del usuario:

  — `hoodie.datasource.hive_sync.jdbcurl` se ha configurado en la URL del servidor de Hive del clúster y ya no es necesario especificarlo. Esto resulta especialmente útil cuando se ejecuta un trabajo en el modo de clúster de Spark, en el que anteriormente había que especificar la IP maestra de Amazon EMR.

  — configuraciones HBase específicas, que son útiles para usar el índice con Hudi. HBase

  — Configuración específica para el proveedor de bloqueos de ZooKeeper, tal como se explica en la sección de control de simultaneidad, que facilita el uso del control de simultaneidad optimista (OCC).
- Se han introducido cambios adicionales para reducir el número de configuraciones que hay que pasar y, en la medida de lo posible, deducir automáticamente:

  — La palabra clave `partitionBy` se puede utilizar para especificar la columna de particiones.

  — Al activar Hive Sync, ya no es obligatorio pasar `HIVE_TABLE_OPT_KEY`, `HIVE_PARTITION_FIELDS_OPT_KEY`, `HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`. Estos valores se pueden deducir del nombre de la tabla de Hudi y del campo de partición.

  — No es obligatorio pasar `KEYGENERATOR_CLASS_OPT_KEY` y se puede deducir de casos más simples de `SimpleKeyGenerator` y `ComplexKeyGenerator`.

Advertencias de Hudi

- Hudi no admite la ejecución vectorizada en las tablas Hive for Merge on Read (MoR) y Bootstrap. Por ejemplo, `count(*)` presenta un error con la tabla en tiempo real de Hudi cuando `hive.vectorized.execution.enabled` se establece en true. Como solución alternativa, puede deshabilitar la lectura vectorizada al establecer `hive.vectorized.execution.enabled` en `false`.
- La compatibilidad con varios escritores no es compatible con la característica de arranque de Hudi.

- Flink Streamer y Flink SQL son características experimentales de esta versión. Se recomienda el uso de estas características en implementaciones de producción.

Cambios, mejoras y problemas resueltos

Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.

- Anteriormente, el reinicio manual del administrador de recursos en un clúster multimaestro provocaba que los daemons en el clúster de Amazon EMR, como ZooKeeper, recargaran todos los nodos perdidos o retirados previamente en el archivo znode de ZooKeeper. Esto provocaba que se superaran los límites predeterminados en ciertas situaciones. Amazon EMR ahora elimina del archivo de ZooKeeper los registros de nodos perdidos o retirados que tengan más de una hora de antigüedad y se han incrementado los límites internos.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.

- Configuración de un clúster para corregir los problemas de rendimiento del servidor YARN Timeline de Apache en las versiones 1 y 1.5

  Las versiones 1 y 1.5 del servidor YARN Timeline de Apache pueden provocar problemas de rendimiento con clústeres de EMR grandes y muy activos, especialmente con `yarn.resourcemanager.system-metrics-publisher.enabled=true`, que es la configuración predeterminada de Amazon EMR. Un servidor YARN Timeline v2 de código abierto resuelve el problema de rendimiento relacionado con la escalabilidad del servidor YARN Timeline.

  Otras alternativas para este problema incluyen:

  - Configuración de yarn.resourcemanager. system-metrics-publisher.enabled=false en yarn-site.xml.
  - La habilitación de la corrección para este problema al crear un clúster, tal y como se describe a continuación.

  Las siguientes versiones de Amazon EMR contienen una corrección para este problema de rendimiento del servidor YARN Timeline.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  Para habilitar la corrección en cualquiera de las versiones de Amazon EMR especificadas anteriormente, defina estas propiedades como `true` en un archivo JSON de configuración que se pasa mediante el parámetro de comando `aws emr create-cluster`: `--configurations file://./configurations.json`. También puede habilitar la corrección mediante la interfaz de usuario de la consola de reconfiguración.

  Ejemplo del contenido del archivo configurations.json:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- Los servidores WebHDFS y HttpFS están deshabilitados de forma predeterminada. Puede volver a habilitar WebHDFS mediante la configuración de Hadoop, `dfs.webhdfs.enabled`. El servidor HttpFS se puede iniciar mediante `sudo systemctl start hadoop-httpfs`.
- HTTPS ahora está habilitado de forma predeterminada para los repositorios de Amazon Linux. Si utiliza una política de VPCE de Amazon S3 para restringir el acceso a buckets específicos, debe agregar el nuevo ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*` del bucket de Amazon Linux a su política (sustituya `$region` por la región en la que se encuentra el punto de conexión). Para obtener más información, consulte este tema en los foros de discusión. AWS Anuncio: Amazon Linux 2 ahora admite la posibilidad de utilizar HTTPS mientras se conecta a repositorios de paquetes.
- Hive: se ha mejorado el rendimiento de las consultas de escritura al permitir el uso de un directorio temporal en HDFS para el último trabajo. Los datos temporales del trabajo final se escriben en HDFS en lugar de Amazon S3 y el rendimiento mejora porque los datos se mueven de HDFS a la ubicación de la tabla final (Amazon S3) en lugar de entre dispositivos Amazon S3.
- Hive: mejora el tiempo de compilación de consultas hasta 2,5 veces con la eliminación de particiones del metaalmacén de Glue.
- De forma predeterminada, cuando Hive UDFs transfiere los elementos integrados al servidor Hive Metastore, solo un subconjunto de los integrados UDFs se pasa a Glue Metastore, ya que Glue solo admite operadores de expresiones limitadas. Si configura `hive.glue.partition.pruning.client=true`, toda la eliminación de particiones se realiza en el cliente. Si configura `hive.glue.partition.pruning.server=true`, entonces toda la eliminación de particiones se realiza en el lado del servidor.

Problemas conocidos

- Las consultas de Hue no funcionan en Amazon EMR 6.4.0 porque el servidor HttpFS de Apache Hadoop está deshabilitado de forma predeterminada. Para usar Hue en Amazon EMR 6.4.0, inicie manualmente el servidor HttpFS en el nodo principal de Amazon EMR mediante `sudo systemctl start hadoop-httpfs` o utilice un paso de Amazon EMR.
- La característica Cuadernos de Amazon EMR que se utiliza con la suplantación de usuarios de Livy no funciona porque HttpFS está deshabilitado de forma predeterminada. En este caso, el cuaderno de EMR no puede conectarse al clúster que tiene habilitada la suplantación de Livy. La solución alternativa consiste en iniciar el servidor HttpFS antes de conectar el cuaderno de EMR al clúster mediante `sudo systemctl start hadoop-httpfs`.
- En la versión 6.4.0 de Amazon EMR, Phoenix no admite el componente de conectores de Phoenix.

- Para utilizar las acciones de Spark con Apache Oozie, debe agregar la siguiente configuración al archivo `workflow.xml` de Oozie. De lo contrario, varias bibliotecas críticas, como Hadoop y EMRFS, no aparecerán en la ruta de clases de los ejecutores de Spark que lance Oozie.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 5.32.0

Las siguientes notas de la versión incluyen información sobre la versión 5.32.0 de Amazon EMR. Los cambios son respecto a la versión 5.31.0.

Fecha de lanzamiento inicial: 8 de enero de 2021

Actualizaciones

- Se ha actualizado el conector de Amazon Glue a la versión 1.14.0
- Se actualizó el SDK de Amazon SageMaker Spark a la versión 1.4.1
- Se actualizó AWS SDK para Java a la versión 1.11.890
- Se ha actualizado el conector de DynamoDB para EMR a la versión 4.16.0

- Se ha actualizado EMRFS a la versión 2.45.0
- Se han actualizado las métricas de análisis de registros de EMR a la versión 1.18.0
- MetricsAndEventsApiGateway Cliente EMR actualizado a la versión 1.5.0
- Se ha actualizado el servidor de registros de EMR a la versión 1.8.0
- Se ha actualizado EMR S3 Dist CP a la versión 2.17.0
- Se ha actualizado EMR Secret Agent a la versión 1.7.0
- Se ha actualizado Flink a la versión 1.11.2
- Se ha actualizado Hadoop a la versión 2.10.1-amzn-0
- Se ha actualizado Hive a la versión 2.3.7-amzn-3
- Se ha actualizado Hue a la versión 4.8.0
- Se ha actualizado MXNet a la versión 1.7.0
- Se ha actualizado OpenCV a la versión 4.4.0
- Se ha actualizado Presto a la versión 0.240.1-amzn-0
- Se ha actualizado Spark a la versión 2.4.7-amzn-0
- Actualizado a la versión TensorFlow 2.3.1

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.

- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Versiones de componentes actualizadas.
- Para obtener una lista de las versiones de los componentes, consulte Acerca de las versiones de Amazon EMR en esta guía.

Nuevas características

- A partir de Amazon EMR 5.32.0 y 6.5.0, el tamaño dinámico del ejecutor para Apache Spark se habilita de forma predeterminada. Para activar o desactivar esta característica, puede utilizar el parámetro de configuración `spark.yarn.heterogeneousExecutors.enabled`.
- Estado de soporte de Instance Metadata Service (IMDS) V2: los componentes de Amazon EMR 5.23.1, 5.27.1 y 5.32 o versiones posteriores se utilizan para todas las llamadas de IMDS. IMDSv2 Para las llamadas de IMDS en el código de su aplicación, puede usar ambos o configurar el IMDS para que se use solo con el IMDSv1 fin de IMDSv2 aumentar la seguridad. IMDSv2 En otras versiones de EMR 5.x, la IMDSv1 desactivación provoca un error al iniciar el clúster.
- A partir de Amazon EMR 5.32.0, puede iniciar un clúster que se integre de forma nativa con Apache Ranger. Apache Ranger es un marco de código abierto para habilitar, supervisar y administrar la seguridad integral de los datos en toda la plataforma Hadoop. Para obtener más información, consulte Apache Ranger. Con la integración nativa, puede utilizar su propio Apache Ranger para aplicar un control de acceso a los datos detallado en Amazon EMR. Consulte Integración de Amazon EMR con Apache Ranger en la Guía de versiones de Amazon EMR.

- La versión 5.32.0 de Amazon EMR es compatible con Amazon EMR en EKS. Para más información sobre cómo empezar a utilizar EMR en EKS, consulte ¿Qué es Amazon EMR en EKS?.
- La versión 5.32.0 de Amazon EMR es compatible con Amazon EMR Studio (versión preliminar). Para más información sobre cómo empezar a usar EMR Studio, consulte Amazon EMR Studio (versión preliminar).
- Políticas gestionadas con ámbito de aplicación: para ajustarse a las prácticas AWS recomendadas, Amazon EMR ha introducido la versión 2 de las políticas gestionadas predeterminadas con ámbito de EMR como sustitutivas de las políticas que quedarán en desuso. Consulte Políticas administradas por Amazon EMR.

Problemas conocidos

- En el caso de los clústeres de subredes privadas de Amazon EMR 6.3.0 y 6.2.0, no puede acceder a la interfaz de usuario web de Ganglia. Aparecerá el error “Acceso denegado (403)”. Otras páginas web UIs, como Spark, Hue, Zeppelin, Livy y JupyterHub Tez, funcionan con normalidad. El acceso a la interfaz de usuario web de Ganglia en los clústeres de subredes públicas también funciona con normalidad. Para resolver este problema, reinicie el servicio httpd en el nodo principal con `sudo systemctl restart httpd`. Este problema se ha corregido en la versión 6.4.0 de Amazon EMR.
- Reducir el límite de «máximo de archivos abiertos» en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
```

```
sudo systemctl daemon-reload
```

- > **Important**
  >
  > Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan automáticamente esas actualizaciones. Para obtener actualizaciones del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.

- Actualmente, la GovCloud región no admite la compatibilidad con consolas para crear una configuración de seguridad que especifique la opción de integración con AWS Ranger. La configuración de seguridad se puede llevar a cabo con la CLI. Consulte Creación de la configuración de seguridad de EMR en la Guía de administración de Amazon EMR.
- Cuando se AtRestEncryption habilita el cifrado HDFS en un clúster que usa Amazon EMR 5.31.0 o 5.32.0, las consultas de Hive producen la siguiente excepción de tiempo de ejecución.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a  b`.

- El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter / (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a  b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 6.2.0

Las siguientes notas de la versión incluyen información sobre la versión 6.2.0 de Amazon EMR. Los cambios son respecto a la versión 6.1.0.

Fecha de lanzamiento inicial: 9 de diciembre de 2020

Fecha de la última actualización: 4 de octubre de 2021

Aplicaciones compatibles

- AWS SDK para Java versión 1.11.828
- emr-record-server versión 1.7.0
- Flink, versión 1.11.2
- Versión de Ganglia 3.7.2
- Hadoop, versión 3.2.1-amzn-1
- HBase versión 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog versión 3.1.2-amzn-0
- Hive, versión 3.1.2-amzn-3
- Hudi versión 0.6.0-amzn-1
- Hue versión 4.8.0
- JupyterHub versión 1.1.0
- Livy versión 0.7.0

- MXNet versión 1.7.0
- Oozie versión 5.2.0
- Versión de Phoenix 5.0.0
- Pig, versión 0.17.0
- Presto, versión 0.238.3-amzn-1
- PrestoSQL, versión 343
- Spark, versión 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow versión 2.3.1
- Zeppelin, versión 0.9.0-preview1
- Versión de Zookeeper 3.4.14
- Conectores y controladores: conector de DynamoDB 4.16.0

Nuevas características

- HBase: Se ha eliminado el cambio de nombre en la fase de confirmación y se ha añadido el seguimiento persistente HFile . Consulte el HFile seguimiento persistente en la guía de versiones de Amazon EMR.
- HBase: Backported Cree una configuración que obligue a almacenar en caché los bloques al compactarlos.
- PrestoDB: mejoras en la eliminación dinámica de particiones. El reordenamiento de uniones basado en reglas funciona con datos no particionados.
- Políticas gestionadas con ámbito de aplicación: para ajustarse a las prácticas AWS recomendadas, Amazon EMR ha introducido la versión 2 de las políticas gestionadas predeterminadas con ámbito de EMR como sustitutivas de las políticas que quedarán en desuso. Consulte Políticas administradas por Amazon EMR.
- Estado de soporte de Instance Metadata Service (IMDS) V2: para Amazon EMR 6.2 o posterior, los componentes de Amazon EMR se IMDSv2 utilizan para todas las llamadas de IMDS. Para las llamadas de IMDS en el código de su aplicación, puede utilizar ambos IMDSv1 o configurar el IMDS para que se utilice únicamente con fines de seguridad adicionales. IMDSv2 IMDSv2 Si la deshabilita IMDSv1 en versiones anteriores de Amazon EMR 6.x, se produce un error al iniciar el clúster.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Spark: mejoras de rendimiento en el tiempo de ejecución de Spark.

Problemas conocidos

- Amazon EMR 6.2 tiene permisos incorrectos configurados en el directorio/etc/cron.d/libinstance-controller-java file in EMR 6.2.0. Permissions on the file are 645 (-rw-r--r-x), when they should be 644 (-rw-r--r--). As a result, Amazon EMR version 6.2 does not log instance-state logs, and the /emr/instance-logs que está vacío. Este problema se ha corregido en la versión 6.3.0 de Amazon EMR y posteriores.

  Para solucionar este problema, ejecute el siguiente script como acción de arranque al iniciar el clúster.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- En el caso de los clústeres de subredes privadas de Amazon EMR 6.2.0 y 6.3.0, no puede acceder a la interfaz de usuario web de Ganglia. Aparecerá el error “Acceso denegado (403)”. Otras páginas web UIs, como Spark, Hue, Zeppelin JupyterHub, Livy y Tez, funcionan con normalidad. El acceso a la interfaz de usuario web de Ganglia en los clústeres de subredes públicas también funciona con normalidad. Para resolver este problema, reinicie el servicio httpd en el nodo principal con `sudo systemctl restart httpd`. Este problema se ha corregido en la versión 6.4.0 de Amazon EMR.
- Hay un problema en Amazon EMR 6.2.0 por el que httpd genera errores continuamente, lo que provoca que Ganglia no esté disponible. Aparece el error “No se puede conectar al servidor”. Para corregir un clúster que ya se está ejecutando con este problema, utilice SSH en el nodo principal del clúster y agregue la línea `Listen 80` al archivo `httpd.conf` ubicado en `/etc/httpd/conf/httpd.conf`. Este problema se ha corregido en la versión 6.3.0 de Amazon EMR.
- HTTPD presenta errores en los clústeres de EMR 6.2.0 cuando se utiliza una configuración de seguridad. Esto hace que la interfaz de usuario de la aplicación web de Ganglia no esté disponible. Para acceder a la interfaz de usuario de la aplicación web de Ganglia, agregue `Listen 80` al archivo `/etc/httpd/conf/httpd.conf` del nodo principal del clúster. Para obtener información acerca de la conexión al clúster, consulte Conectarse al nodo principal mediante SSH.

  Cuadernos de EMR tampoco puede establecer una conexión con los clústeres de EMR 6.2.0 cuando se utiliza una configuración de seguridad. El cuaderno no podrá enumerar los kernels ni enviar trabajos a Spark. En su lugar, se recomienda que utilice Cuadernos de EMR con otra versión de Amazon EMR.

- Reducir el límite de «máximo de archivos abiertos» en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

  Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

  Establecimiento explícito de un ulimit desde la línea de comandos

  1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

     ```
     LimitNOFILE=65536

     LimitNPROC=65536
     ```

  2. Reinicie InstanceController

     ```
     $ sudo systemctl daemon-reload

     $ sudo systemctl restart instance-controller
     ```

  Establecimiento de un ulimit mediante una acción de arranque (BA)

  También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon EMR 6.1.0 y 6.2.0 incluyen un problema de rendimiento que puede afectar gravemente a todas las operaciones de inserción, upsert y eliminación de Hudi. Si piensa utilizar Hudi con Amazon EMR 6.1.0 o 6.2.0, póngase en AWS contacto con el servicio de asistencia para obtener un RPM de Hudi parcheado.

- **Important**

  Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan automáticamente esas actualizaciones. Para obtener actualizaciones

del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.

- Los artefactos de Maven de Amazon EMR 6.2.0 no están publicados. Se publicarán en una versión futura de Amazon EMR.
- El HFile seguimiento persistente mediante la tabla del sistema de HBase archivos de almacenamiento no admite la función de replicación HBase regional. Para obtener más información sobre HBase la replicación regional, consulte Lecturas de alta disponibilidad consistentes en el cronograma.
- Diferencias entre las versiones de agrupación en buckets de Amazon EMR 6.x y EMR 5.x Hive

  EMR 5.x utiliza OOS Apache Hive 2, mientras que en EMR 6.x se utiliza OOS Apache Hive 3. Hive2 de código abierto usa la versión 1 de agrupación en buckets, mientras que Hive3 de código abierto usa la versión 2 de agrupación en buckets. Esta diferencia de versión de agrupación en buckets entre Hive 2 (EMR 5.x) y Hive 3 (EMR 6.x) significa que el hash de agrupación en buckets de Hive funciona de manera diferente. Vea el ejemplo siguiente.

  La siguiente tabla es un ejemplo creado en EMR 6.x y EMR 5.x, respectivamente.

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

  Se insertan los mismos datos en EMR 6.x y EMR 5.x.

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

  Al comprobar la ubicación de S3, se muestra que el nombre del archivo de agrupación en buckets es diferente, ya que la función de hash es diferente entre EMR 6.x (Hive 3) y EMR 5.x (Hive 2).

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
```

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

También puede ver la diferencia de versión mediante la ejecución del siguiente comando en la CLI de Hive en EMR 6.x. Tenga en cuenta que devuelve la versión 2 de la agrupación en buckets.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

    > **Note**
    >
    > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede

configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 5.31.0

Las siguientes notas de la versión incluyen información sobre la versión 5.31.0 de Amazon EMR. Los cambios son respecto a la versión 5.30.1.

Fecha de lanzamiento inicial: 9 de octubre de 2020

Fecha de la última actualización: 15 de octubre de 2020

Actualizaciones

- Se ha actualizado el conector de Amazon Glue a la versión 1.13.0
- Se actualizó el SDK de Amazon SageMaker Spark a la versión 1.4.0
- Se ha actualizado el conector Amazon Kinesis a la versión 3.5.9
- Se actualizó AWS SDK para Java a la versión 1.11.852
- Se ha actualizado Bigtop-tomcat a la versión 8.5.56

- Se ha actualizado EMRFS a la versión 2.43.0
- MetricsAndEventsApiGateway Cliente EMR actualizado a la versión 1.4.0
- Se ha actualizado EMR S3 Dist CP a la versión 2.15.0
- Se ha actualizado EMR S3 Select a la versión 1.6.0
- Se ha actualizado Flink a la versión 1.11.0
- Se ha actualizado Hadoop a la versión 2.10.0
- Se ha actualizado Hive a la versión 2.3.7
- Se ha actualizado Hudi a la versión 0.6.0
- Se ha actualizado Hue a la versión 4.7.1
- Actualizado a la versión JupyterHub 1.1.0
- Se ha actualizado MXNet a la versión 1.6.0
- Se ha actualizado OpenCV a la versión 4.3.0
- Se ha actualizado Presto a la versión 0.238.3
- Actualizado TensorFlow a la versión 2.1.0

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.

- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Las estadísticas de columnas de Hive son compatibles con las versiones 5.31.0 y posteriores de Amazon EMR.
- Versiones de componentes actualizadas.
- Compatibilidad con EMRFS S3EC V2 en Amazon EMR 5.31.0. En las versiones 1.11.837 y posteriores del SDK para Java de S3, se ha introducido la versión 2 del cliente de cifrado (S3EC V2) con varias mejoras de seguridad. Para obtener más información, consulte los siguientes temas:
  - Publicación en el blog de S3: Updates to the Amazon S3 encryption client.
  - AWS SDK para Java Guía para desarrolladores: migre los clientes de cifrado y descifrado a la versión 2.
  - Guía de administración de EMR: Cifrado del cliente de Amazon S3.

  La versión 1 del cliente de cifrado todavía está disponible en el SDK por motivos de compatibilidad con versiones anteriores.

Nuevas características

- Reducir el límite de «máximo de archivos abiertos» en los más antiguos AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora

incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
```

```
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Con Amazon EMR 5.31.0, puede lanzar un clúster que se integre con Lake Formation. Esta integración proporciona un filtrado de datos detallado a nivel de columna para las bases de datos y tablas del catálogo de datos de Glue. AWS También permite un inicio de sesión único federado en cuadernos de EMR o Apache Zeppelin desde un sistema de identidades empresariales. Para más información, consulte Integración de Amazon EMR con AWS Lake Formation en la Guía de administración de Amazon EMR.

  Amazon EMR with Lake Formation está disponible actualmente en 16 AWS regiones: EE.UU. Este (Ohio y Norte de Virginia), EE.UU. Oeste (Norte de California y Oregón), Asia Pacífico (Bombay, Seúl, Singapur, Sídney y Tokio), Canadá (Centro), Europa (Fráncfort, Irlanda, Londres, París y Estocolmo) y Sudamérica (São Paulo).

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

  ```
  kinit -kt <keytab_file> <principal>
  ```

  Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Cuando se AtRestEncryption habilita el cifrado HDFS en un clúster que usa Amazon EMR 5.31.0 o 5.32.0, las consultas de Hive producen la siguiente excepción de tiempo de ejecución.

  ```
  TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
   attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
   java.lang.RuntimeException: Hive Runtime Error while closing
   operators: java.io.IOException: java.util.ServiceConfigurationError:
   org.apache.hadoop.security.token.TokenIdentifier: Provider
   org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
  ```

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a  b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a  b` entra en esta categoría. Tenga en cuenta que hay otros

14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 6.1.0

Las siguientes notas de la versión incluyen información sobre la versión 6.1.0 de Amazon EMR. Los cambios son respecto a la versión 6.0.0.

Fecha de lanzamiento inicial: 4 de septiembre de 2020

Fecha de la última actualización: 15 de octubre de 2020

Aplicaciones compatibles

- AWS SDK para Java versión 1.11.828
- Flink, versión 1.11.0
- Versión de Ganglia 3.7.2
- Hadoop, versión 3.2.1-amzn-1
- HBase versión 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog versión 3.1.2-amzn-0
- Hive, versión 3.1.2-amzn-1
- Hudi versión 0.5.2-incubating
- Hue versión 4.7.1
- JupyterHub versión 1.1.0
- Livy versión 0.7.0
- MXNet versión 1.6.0
- Oozie versión 5.2.0
- Versión de Phoenix 5.0.0
- Presto, versión 0.232

- PrestoSQL, versión 338
- Spark, versión 3.0.0-amzn-0
- TensorFlow versión 2.1.0
- Zeppelin, versión 0.9.0-preview1
- Versión de Zookeeper 3.4.14
- Conectores y controladores: DynamoDB Connector 4.14.0

Nuevas características

- Los tipos de instancias de ARM se admiten a partir de la versión 5.30.0 de Amazon EMR y la versión 6.1.0 de Amazon EMR.
- Los tipos de instancia de uso general m6g se admiten a partir de las versiones 6.1.0 y 5.30.0 de Amazon EMR. Para más información, consulte Tipos de instancias admitidas en la Guía de administración de Amazon EMR.
- La función de grupo de EC2 ubicación se admite a partir de la versión 5.23.0 de Amazon EMR como opción para varios clústeres de nodos principales. Actualmente, la característica de grupos de ubicación solo admite los tipos de nodos principales y la estrategia `SPREAD` se aplica a esos nodos. La estrategia `SPREAD` coloca un pequeño grupo de instancias en un hardware subyacente independiente para evitar la pérdida de varios nodos principales en caso de que se produzca un error de hardware. Para obtener más información, consulte Integración de EMR con EC2 Placement Group en la Guía de administración de Amazon EMR.
- Escalado administrado: con la versión 6.1.0 de Amazon EMR, puede habilitar el escalado administrado de Amazon EMR para aumentar o disminuir automáticamente el número de instancias o unidades del clúster en función de la carga de trabajo. Amazon EMR evalúa continuamente las métricas del clúster para tomar decisiones de escalado que optimicen los clústeres en cuanto al costo y la velocidad. El escalado administrado también está disponible en la versión 5.30.0 y posteriores de Amazon EMR, excepto 6.0.0. Para más información, consulte Recursos de escalado de clúster en la Guía de administración de Amazon EMR.
- La versión 338 de PrestoSQL es compatible con EMR 6.1.0. Para más información, consulte Presto.
  - PrestoSQL solo es compatible con EMR 6.1.0 y versiones posteriores, no con EMR 6.0.0 ni EMR 5.x.
  - El nombre de la aplicación, `Presto`, se sigue utilizando para instalar PrestoDB en los clústeres. Para instalar PrestoSQL en los clústeres, utilice el nombre de la aplicación `PrestoSQL`.

  - Puede instalar PrestoDB o PrestoSQL, pero no puede instalar ambos en un solo clúster. Si se especifican PrestoDB y PrestoSQL al intentar crear un clúster, se produce un error tanto en la validación como en la solicitud de creación del clúster.
  - PrestoSQL es compatible con clústeres de maestro único o multimaestro. En los clústeres multimaestro, se necesita un metaalmacén de Hive externo para ejecutar PrestoSQL o PrestoDB. Consulte Aplicaciones admitidas en un clúster de EMR con varios nodos principales.
- Soporte de autenticación automática de ECR en Apache Hadoop y Apache Spark con Docker: los usuarios de Spark pueden utilizar imágenes de Docker de Docker Hub y Amazon Elastic Container Registry (Amazon ECR) para definir las dependencias del entorno y la biblioteca.

  Configuración de Docker y Ejecución de aplicaciones de Spark con Docker mediante Amazon EMR 6.x..
- EMR admite transacciones ACID de Apache Hive: Amazon EMR 6.1.0 agrega compatibilidad con las transacciones ACID de Hive para que cumpla con las propiedades ACID de una base de datos. Con esta característica, puede ejecutar operaciones `INSERT, UPDATE, DELETE,` y `MERGE` en tablas administradas por Hive con datos de Amazon Simple Storage Service (Amazon S3). Se trata de una característica clave para casos de uso como la ingesta de contenido en streaming, la reformulación de datos, las actualizaciones masivas mediante MERGE y los cambios lentos de dimensiones. Para más información, incluidos ejemplos de configuración y casos de uso, consulte Amazon EMR admite transacciones ACID de Apache Hive.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.

- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Apache Flink no es compatible con EMR 6.0.0, pero sí con EMR 6.1.0 con Flink 1.11.0. Esta es la primera versión de Flink que es compatible oficialmente con Hadoop 3. Consulte Apache Flink 1.11.0 Release Announcement.
- Se ha eliminado Ganglia de los paquetes predeterminados de EMR 6.1.0.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de

archivos abiertos", una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para "Máximo de archivos abiertos" provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
```

```
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- > **Important**
  >
  > Amazon EMR 6.1.0 y 6.2.0 incluyen un problema de rendimiento que puede afectar gravemente a todas las operaciones de inserción, upsert y eliminación de Hudi. Si piensa utilizar Hudi con Amazon EMR 6.1.0 o 6.2.0, póngase en AWS contacto con el servicio de asistencia para obtener un RPM de Hudi parcheado.

- Si establece una configuración de recolección de basura personalizada con `spark.driver.extraJavaOptions` y`spark.executor.extraJavaOptions`, esto provocará un error de driver/executor inicio con EMR 6.1 debido a un conflicto en la configuración de recolección de basura. En su lugar, con la versión 6.1.0 de EMR, debe especificar una configuración personalizada de recopilación de elementos no utilizados de Spark para los controladores y ejecutores con las propiedades `spark.driver.defaultJavaOptions` y `spark.executor.defaultJavaOptions`. Obtenga más información en Apache Spark Runtime Environment y Configuración de la recopilación de elementos no utilizados de Spark en Amazon EMR 6.1.0.
- Al utilizar Pig con Oozie (y dentro de Hue, ya que Hue utiliza las acciones de Oozie para ejecutar scripts de Pig), se genera un error que indica que no se puede cargar una biblioteca lzo nativa. Este mensaje de error es informativo y no impide que Pig se ejecute.
- Soporte de simultaneidad de Hudi: actualmente, Hudi no admite escrituras simultáneas en una sola tabla de Hudi. Además, Hudi revierte los cambios que estén realizando los escritores en curso antes de permitir que comience un nuevo escritor. Las escrituras simultáneas pueden interferir con este mecanismo e introducir condiciones de carrera, lo que puede provocar la corrupción de los datos. Debe asegurarse de que, como parte de su flujo de trabajo de procesamiento de datos, solo haya un escritor de Hudi trabajando en una tabla de Hudi a la vez. Hudi admite que varios lectores simultáneos trabajen en la misma tabla de Hudi.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el

clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Hay un problema en Amazon EMR 6.1.0 que afecta a los clústeres que ejecutan Presto. Tras un periodo de tiempo prolongado (días), el clúster puede generar errores como “su: no se puede ejecutar /bin/bash: recurso temporalmente no disponible” o “error de la solicitud del intérprete de comandos en el canal 0”. Este problema se debe a un proceso interno de Amazon EMR (InstanceController) que genera demasiados procesos ligeros (LWP), lo que, finalmente, hace que el usuario de Hadoop supere su límite de nproc. Esto impide que el usuario abra procesos adicionales. La solución para este problema consiste en actualizar a EMR 6.2.0.

Versión 6.0.0

Las siguientes notas de la versión incluyen información sobre la versión 6.0.0 de Amazon EMR.

Fecha de lanzamiento inicial: 10 de marzo de 2020

Aplicaciones compatibles

- AWS SDK para Java versión 1.11.711
- Versión de Ganglia 3.7.2
- Versión de Hadoop 3.2.1
- HBase versión 2.2.3
- HCatalog versión 3.1.2
- Versión de Hive 3.1.2
- Versión de Hudi 0.5.0-incubating
- Versión de Hue 4.4.0
- JupyterHub versión 1.0.0
- Versión de Livy 0.6.0
- MXNet versión 1.5.1
- Versión de Oozie 5.1.0
- Versión de Phoenix 5.0.0
- Versión de Presto 0.230
- Versión de Spark 2.4.4
- TensorFlow versión 1.14.0
- Versión de Zeppelin 0.9.0-SNAPSHOT
- Versión de Zookeeper 3.4.14
- Conectores y controladores: DynamoDB Connector 4.14.0

> **Note**
>
> Flink, Sqoop, Pig y Mahout no están disponibles en la versión 6.0.0 de Amazon EMR.

Nuevas características

- Soporte de tiempo de ejecución para YARN Docker: las aplicaciones de YARN, como los trabajos de Spark, ahora pueden ejecutarse en el contexto de un contenedor de Docker. Esto le permite definir fácilmente dependencias en una imagen de Docker sin necesidad de instalar bibliotecas

personalizadas en su clúster de Amazon EMR. Para más información, consulte Configurar la integración de Docker y Ejecutar aplicaciones de Spark con Docker mediante Amazon EMR 6.0.0.

- Soporte LLAP de Hive: Hive ahora admite el modo de ejecución LLAP para mejorar el rendimiento de las consultas. Para obtener más información, consulte Uso de Hive LLAP.

Cambios, mejoras y problemas resueltos

- Esta es una versión para solucionar problemas con Amazon EMR Scaling cuando no logra up/scale reducir correctamente un clúster o provoca errores en las aplicaciones.
- Se solucionó un problema por el que las solicitudes de escalado de un clúster grande y muy utilizado fallaban cuando los daemons en el clúster de Amazon EMR ejecutaban actividades de comprobación de estado, como recopilar el estado del nodo de YARN y el estado del nodo de HDFS. Esto se debía a que los daemons en el clúster no podían comunicar los datos del estado de un nodo a los componentes internos de Amazon EMR.
- Se han mejorado los daemons de EMR en el clúster para realizar un seguimiento correcto de los estados de los nodos cuando se reutilizan las direcciones IP a fin de mejorar la fiabilidad durante las operaciones de escalado.
- SPARK-29683. Se ha corregido un error que provocaba que se produjeran errores en los trabajos durante la reducción vertical del clúster, ya que Spark daba por sentado que todos los nodos disponibles estaban en la lista de denegados.
- YARN-9011. Se ha corregido un error que provocaba que se produjeran errores en los trabajos debido a una condición de carrera durante la retirada de YARN cuando el clúster intentaba escalarse o reducirse verticalmente.
- Se ha solucionado el problema que provocaba errores en los pasos o trabajos durante el escalado del clúster al garantizar que los estados de los nodos sean siempre coherentes entre los daemons en el clúster de Amazon EMR y YARN o HDFS.
- Se ha solucionado un problema por el que las operaciones de clúster, como la reducción vertical y el envío escalonado, fallaban en los clústeres de Amazon EMR habilitados con la autenticación de Kerberos. Esto se debió a que el daemon del clúster de Amazon EMR no renovó el ticket de Kerberos, que es necesario para comunicarse de forma segura con la HDFS/YARN ejecución en el nodo principal.
- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.

- Amazon Linux
  - Amazon Linux 2 es el sistema operativo de la serie de versiones 6.x de EMR.
  - Para la administración de servicios se utiliza `systemd` en lugar de `upstart` que se utilizaba en Amazon Linux 1.
- Kit de desarrollo de Java (JDK)
  - Corretto JDK 8 es el JDK predeterminado para la serie de versiones de EMR 6.x.
- Scala
  - Scala 2.12 se utiliza con Apache Spark y Apache Livy.
- Python 3
  - Python 3 es ahora la versión predeterminada de Python en EMR.
- Etiquetas de nodo YARN
  - A partir de la serie de versiones 6.x de Amazon EMR, la característica de etiquetas de nodo YARN está desactivada de forma predeterminada. De forma predeterminada, los procesos maestros de la aplicación se pueden ejecutar tanto en nodos principales como en nodos de tareas. Puede habilitar la función de etiquetas de nodo YARN configurando las siguientes propiedades: `yarn.node-labels.enabled` y `yarn.node-labels.am.default-node-label-expression`. Para más información, consulte Comprensión de los nodos principales, básicos y de tareas.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de

4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
```

```
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- El shell interactivo de Spark, que incluye PySpark SparkR y spark-shell, no admite el uso de Docker con bibliotecas adicionales.
- Para utilizar Python 3 con la versión 6.0.0 de Amazon EMR, debe agregar PATH a `yarn.nodemanager.env-whitelist`.
- La funcionalidad Live Long and Process (LLAP) no se admite cuando se utiliza el catálogo de datos de AWS Glue como metaalmacén de Hive.
- Al utilizar Amazon EMR 6.0.0 con la integración de Spark y Docker, debe configurar las instancias del clúster con el mismo tipo de instancia y la misma cantidad de volúmenes de EBS para evitar errores al enviar un trabajo de Spark con el tiempo de ejecución de Docker.
- En Amazon EMR 6.0.0, el modo de almacenamiento de Amazon HBase S3 se ve afectado por el problema del HBASE-24286. HBase el maestro no se puede inicializar cuando se crea el clúster con los datos de S3 existentes.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.30.1

Las siguientes notas de la versión incluyen información sobre la versión 5.30.1 de Amazon EMR. Los cambios son respecto a la versión 5.30.0.

Fecha de lanzamiento inicial: 30 de junio de 2020

Fecha de la última actualización: 24 de agosto de 2020

Cambios, mejoras y problemas resueltos

- Las versiones más recientes de Amazon EMR solucionan el problema con un límite de «máximo de archivos abiertos» inferior al anterior en AL2 Amazon EMR. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración “Máximo de archivos abiertos” más alta.
- Se ha solucionado el problema que provocaba que el proceso del controlador de instancias generaba un número infinito de procesos.
- Se ha solucionado el problema que provocaba que Hue no pudiera ejecutar una consulta de Hive, mostrara el mensaje “database is locked” e impidiera la ejecución de consultas.
- Se ha corregido un problema de Spark que permitía ejecutar más tareas simultáneamente en el clúster de EMR.
- Se ha corregido un error en el cuaderno de Jupyter que provocaba el “too many files open” en el servidor de Jupyter.
- Se ha corregido un problema con las horas de inicio de los clústeres.

Nuevas características

- Las interfaces de aplicación persistentes de la IU de Tez y el servidor YARN Timeline están disponibles con las versiones 6.x de Amazon EMR y la versión 5.30.1 y posteriores de EMR. El

acceso mediante un enlace al historial de aplicaciones persistente con un solo clic le permite acceder rápidamente al historial de trabajos sin necesidad de configurar un proxy web a través de una conexión SSH. Los registros de los clústeres activos y terminados están disponibles durante los 30 días posteriores a la finalización de la aplicación. Para más información, consulte Ver interfaces de usuario de aplicaciones persistentes en la Guía de administración de Amazon EMR.

- La ejecución de cuadernos EMR APIs está disponible para ejecutar cuadernos EMR mediante un script o una línea de comandos. La capacidad de iniciar, detener, enumerar y describir las ejecuciones de cuadernos EMR sin la AWS consola le permite controlar mediante programación un cuaderno EMR. Al utilizar una celda de cuaderno parametrizada, puede pasar diferentes valores de parámetros a un cuaderno sin tener que crear una copia del cuaderno para cada nuevo conjunto de valores de parámetros. Consulte Acciones de la API de EMR. Para ver un código de ejemplo, consulte Ejemplos de comandos para ejecutar Cuadernos de EMR mediante programación.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

  Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- EMR Notebooks

La característica que le permite instalar kernels y bibliotecas de Python adicionales en el nodo principal del clúster está deshabilitada de forma predeterminada en la versión 5.30.1 de EMR. Para más información sobre esta característica, consulte Instalación de kernels y bibliotecas de Python en el nodo principal del clúster.

Para habilitar la característica, haga lo siguiente:

1. Asegúrese de que la política de permisos adjunta al rol de servicio de Cuadernos de EMR permita la siguiente acción:

   ```
   elasticmapreduce:ListSteps
   ```

   Para obtener más información, consulte Rol de servicio para cuadernos de EMR.

2. Utilice el AWS CLI para ejecutar un paso en el clúster que configura los EMR Notebooks, como se muestra en el siguiente ejemplo. *us-east-1*Sustitúyalo por la región en la que reside el clúster. Para obtener más información, consulte Adición de pasos a un clúster mediante la AWS CLI.

   ```
   aws emr add-steps  --cluster-id MyClusterID --steps
    Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
   east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
   awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
   ```

- Escalado administrado

  Las operaciones de escalado administrado en los clústeres 5.30.0 y 5.30.1 sin Presto instalado pueden provocar errores en las aplicaciones o provocar que un grupo de instancias o una flota de instancias uniformes permanezcan en estado ARRESTED, especialmente cuando una operación de reducción vertical va seguida inmediatamente de una operación de escalado vertical.

  Como solución alternativa, elija Presto como aplicación para instalar cuando cree un clúster con las versiones 5.30.0 y 5.30.1 de Amazon EMR, incluso si su trabajo no requiere Presto.

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del

periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 5.30.0

Las siguientes notas de la versión incluyen información sobre la versión 5.30.0 de Amazon EMR. Los cambios son respecto a la versión 5.29.0.

Fecha de lanzamiento inicial: 13 de mayo de 2020

Fecha de la última actualización: 25 de junio de 2020

Actualizaciones

- Se actualizó AWS SDK para Java a la versión 1.11.759
- Se actualizó el SDK de Amazon SageMaker Spark a la versión 1.3.0
- Se actualizó el servidor de registros de EMR a la versión 1.6.0
- Se ha actualizado Flink a la versión 1.10.0
- Se actualizó Ganglia a la versión 3.7.2
- Se actualizó HBase a la versión 1.4.13
- Se actualizó Hudi a la versión 0.5.2-incubating
- Se actualizó Hue a la versión 4.6.0
- Actualizado a la versión JupyterHub 1.1.0
- Se actualizó Livy a la versión 0.7.0-incubating
- Se actualizó Oozie a la versión 5.2.0
- Se actualizó Presto a la versión 0.232
- Se actualizó Spark a la versión 2.4.5
- Conectores y controladores actualizados: Amazon Glue Connector 1.12.0; Amazon Kinesis Connector 3.5.0; EMR DynamoDB Connector 4.14.0

Nuevas características

- Cuadernos de EMR: cuando se utiliza con clústeres de EMR creados con la versión 5.30.0, los kernels de cuadernos de EMR se ejecutan en un clúster. Esto mejora el rendimiento del bloc de

notas y le permite instalar y personalizar kernels. También puede instalar bibliotecas de Python en el nodo principal del clúster. Para obtener más información, consulte Instalación y uso de kernels y bibliotecas en la Guía de administración de EMR.

- Escalado administrado: con la versión 5.30.0 y posteriores de Amazon EMR, puede habilitar el escalado administrado de EMR para aumentar o disminuir automáticamente el número de instancias o unidades del clúster en función de la carga de trabajo. Amazon EMR evalúa continuamente las métricas del clúster para tomar decisiones de escalado que optimicen los clústeres en cuanto al costo y la velocidad. Para más información, consulte Recursos de escalado de clúster en la Guía de administración de Amazon EMR.
- Cifrar los archivos de registro almacenados en Amazon S3: con Amazon EMR versión 5.30.0 y versiones posteriores, puede cifrar los archivos de registro almacenados en Amazon S3 con una clave gestionada por el cliente. AWS KMS Para más información, consulte Consultar archivos de registros almacenados en Amazon S3 en la Guía de administración de Amazon EMR.
- Compatibilidad con Amazon Linux 2: en la versión 5.30.0 y posteriores de EMR, EMR utiliza el sistema operativo Amazon Linux 2. La nueva personalización AMIs (Amazon Machine Image) debe estar basada en la AMI de Amazon Linux 2. Para obtener más información, consulte Uso de una AMI personalizada.
- Escalado automático estable de Presto: los clústeres de EMR que utilizan la versión 5.30.0 se pueden configurar con un periodo de espera de escalado automático que da tiempo a las tareas de Presto para terminar de ejecutarse antes de que su nodo se retire. Para obtener más información, consulte Uso del escalado automático de Presto con la retirada estable.
- Creación de instancias de flota con una nueva opción de estrategia de asignación: hay una nueva opción de estrategia de asignación disponible en la versión 5.12.1 y posteriores de EMR. Ofrece un aprovisionamiento de clústeres más rápido, una asignación de spot más precisa y menos interrupciones de las instancias de spot. Se requieren actualizaciones de los roles de servicio de EMR no predeterminados. Consulte Configurar flotas de instancias.
- Comandos sudo systemctl stop y sudo systemctl start: en la versión 5.30.0 y posteriores de EMR, que utilizan el sistema operativo Amazon Linux 2, EMR utiliza los comandos `sudo systemctl stop` y `sudo systemctl start` para reiniciar los servicios. Para más información, consulte ¿Cómo reinicio un servicio en Amazon EMR?.

Cambios, mejoras y problemas resueltos

- La versión 5.30.0 de EMR no instala Ganglia de forma predeterminada. Puede seleccionar explícitamente Ganglia para instalar cuando cree un clúster.

- Optimizaciones de rendimiento de Spark.
- Optimizaciones de rendimiento de Presto.
- Python 3 es la opción predeterminada para la versión 5.30.0 y posteriores de Amazon EMR.
- El grupo de seguridad administrado predeterminado para el acceso al servicio en subredes privadas se ha actualizado con nuevas reglas. Si utiliza un grupo de seguridad personalizado para el acceso al servicio, debe incluir las mismas reglas que el grupo de seguridad administrado predeterminado. Para más información, consulte Grupo de seguridad administrado por Amazon EMR para el acceso al servicio (Subredes privadas). Si utiliza un rol de servicio personalizado para Amazon EMR, debe conceder permiso a `ec2:describeSecurityGroups` para que EMR pueda validar si los grupos de seguridad se crean correctamente. Si utiliza `EMR_DefaultRole`, este permiso ya está incluido en la política administrada predeterminada.

Problemas conocidos

- Reducir el límite máximo de archivos abiertos en las versiones anteriores AL2 [corregido en las versiones más recientes]. Versiones de Amazon EMR: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0 y emr-6.2.0 se basan en versiones anteriores de Amazon Linux 2 (), que AL2 tienen una configuración de límite inferior para el «Máximo número de archivos abiertos» cuando los clústeres de Amazon EMR se crean con la AMI predeterminada. Las versiones 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 y posteriores de Amazon EMR ahora incluyen una corrección permanente con una configuración más alta de “Máximo de archivos abiertos”. Las versiones con el límite inferior de archivos abiertos provocan el error “Demasiados archivos abiertos” al enviar el trabajo de Spark. En las versiones afectadas, la AMI predeterminada de Amazon EMR tiene una configuración de ulimit predeterminada de 4096 para “Máximo de archivos abiertos”, una cantidad inferior al límite de 65 536 archivos de la AMI de Amazon Linux 2 más reciente. La configuración de ulimit inferior para “Máximo de archivos abiertos” provoca un fallo en el trabajo de Spark cuando el controlador y el ejecutor de Spark intentan abrir más de 4096 archivos. Para solucionar el problema, Amazon EMR tiene un script de acciones de arranque (BA) que ajusta la configuración de ulimit al crear el clúster.

  Si utiliza una versión anterior de Amazon EMR que no tiene una corrección permanente para este problema, la siguiente solución alternativa le permite establecer explícitamente el ulimit del controlador de instancias en un máximo de 65 536 archivos.

Establecimiento explícito de un ulimit desde la línea de comandos

1. Edite `/etc/systemd/system/instance-controller.service` para agregar los siguientes parámetros a la sección de servicio.

   ```
   LimitNOFILE=65536

   LimitNPROC=65536
   ```

2. Reinicie InstanceController

   ```
   $ sudo systemctl daemon-reload

   $ sudo systemctl restart instance-controller
   ```

Establecimiento de un ulimit mediante una acción de arranque (BA)

También puede usar un script de acciones de arranque (BA) para configurar el ulimit del controlador de instancias en 65 536 archivos al crear el clúster.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Escalado administrado

Las operaciones de escalado administrado en los clústeres 5.30.0 y 5.30.1 sin Presto instalado pueden provocar errores en las aplicaciones o provocar que un grupo de instancias o una flota de instancias uniformes permanezcan en estado `ARRESTED`, especialmente cuando una operación de reducción vertical va seguida inmediatamente de una operación de escalado vertical.

Como solución alternativa, elija Presto como aplicación para instalar cuando cree un clúster con las versiones 5.30.0 y 5.30.1 de Amazon EMR, incluso si su trabajo no requiere Presto.

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

- El motor de base de datos predeterminado para Hue 4.6.0 es SQLite, lo que provoca problemas al intentar utilizar Hue con una base de datos externa. Para solucionar este problema, defina `engine`

en su clasificación de configuración `hue-ini` como `mysql`. Este problema se ha corregido en la versión 5.30.1 de Amazon EMR.

- Cuando utiliza Spark con el formato de ubicación de particiones de Hive para leer datos en Amazon S3 y ejecuta Spark en las versiones 5.30.0 a 5.36.0 y 6.2.0 a 6.9.0 de Amazon EMR, es posible que se produzca un problema que impida que el clúster lea los datos correctamente. Esto puede ocurrir si las particiones tienen todas las características siguientes:
  - Se analizan dos o más particiones de la misma tabla.
  - Al menos una ruta de directorio de particiones es un prefijo de al menos otra ruta de directorio de particiones; por ejemplo, `s3://bucket/table/p=a` es un prefijo de `s3://bucket/table/p=a b`.
  - El primer carácter que sigue al prefijo del otro directorio de particiones tiene un valor UTF-8 inferior al carácter `/` (U+002F). Por ejemplo, el carácter de espacio (U+0020) que aparece entre a y b en `s3://bucket/table/p=a b` entra en esta categoría. Tenga en cuenta que hay otros 14 caracteres que no son de control: `!"#$%&'()*+,-`. Para más información, consulte UTF-8 encoding table and Unicode characters.

  Como solución alternativa a este problema, defina la configuración `spark.sql.sources.fastS3PartitionDiscovery.enabled` como `false` en la clasificación `spark-defaults`.

Versión 5.29.0

Las siguientes notas de la versión incluyen información sobre la versión 5.29.0 de Amazon EMR. Los cambios son respecto a la versión 5.28.1.

Fecha de lanzamiento inicial: 17 de enero de 2020

Actualizaciones

- Se actualizó AWS SDK para Java a la versión 1.11.682
- Actualización de Hive a la versión 2.3.6
- Se ha actualizado Flink a la versión 1.9.1
- Se ha actualizado EMRFS a la versión 2.38.0
- Se ha actualizado el conector EMR DynamoDB a la versión 4.13.0

Cambios, mejoras y problemas resueltos

- Spark
  - Optimizaciones de rendimiento de Spark.
- EMRFS
  - La guía de administración actualiza la configuración predeterminada de emrfs-site.xml para una visualización coherente.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede

configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.28.1

Las siguientes notas de la versión incluyen información sobre la versión 5.28.1 de Amazon EMR. Los cambios son respecto a la versión 5.28.0.

Fecha de lanzamiento inicial: 10 de enero de 2020

Cambios, mejoras y problemas resueltos

- Spark
  - Se han corregido problemas de compatibilidad con Spark.
- CloudWatch Métricas
  - Se corrigió la publicación de Amazon CloudWatch Metrics en un clúster de EMR con varios nodos principales.
- Mensaje de registro deshabilitado
  - Se ha deshabilitado el mensaje de registro falso: “... usando una versión anterior (<4.5.8) del cliente http de Apache”.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

- SSH como usuario de hadoop con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de hadoop.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en /etc/hadoop.keytab y la entidad principal tiene la forma de hadoop/<hostname>@<REALM>.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.28.0

Las siguientes notas de la versión incluyen información sobre la versión 5.28.0 de Amazon EMR. Los cambios son respecto a la versión 5.27.0.

Fecha de lanzamiento inicial: 12 de noviembre de 2019

Actualizaciones

- Actualización de Flink a la versión 1.9.0
- Actualización de Hive a la versión 2.3.6
- Se actualizó MXNet a la versión 1.5.1
- Actualización de Phoenix a la versión 4.14.3
- Actualización de Presto a la versión 0.227
- Actualización de Zeppelin a la versión 0.8.2

Nuevas características

- Apache Hudi ahora está disponible para que Amazon EMR lo instale al crear un clúster. Para obtener más información, consulte Hudi.

- (25 de noviembre de 2019) Ahora puede ejecutar varios pasos en paralelo para mejorar la utilización del clúster y ahorrar costos. También puede cancelar los pasos pendientes y los pasos que estén en ejecución. Para obtener más información, consulte Cómo trabajar con Steps utilizando la consola AWS CLI and.
- (3 de diciembre de 2019) Ahora puede crear y ejecutar clústeres de EMR en ellos. AWS Outposts AWS Outposts habilita AWS los servicios, la infraestructura y los modelos operativos nativos en las instalaciones locales. En AWS Outposts los entornos, puede usar las mismas AWS APIs herramientas e infraestructura que usa en la AWS nube. Para obtener más información, consulte Clústeres de EMR en. AWS Outposts
- (11 de marzo de 2020) A partir de la versión 5.28.0 de Amazon EMR, puede crear y ejecutar clústeres de Amazon EMR en una subred de Zonas AWS locales como una extensión lógica de una región que admite Zonas locales. AWS Una zona local permite que las funciones de Amazon EMR y un subconjunto de AWS servicios, como los servicios de procesamiento y almacenamiento, se ubiquen más cerca de los usuarios, lo que proporciona un acceso de muy baja latencia a las aplicaciones que se ejecutan localmente. Para obtener una lista de las zonas locales disponibles, consulte Zonas locales de AWS. Para obtener información sobre el acceso a las zonas AWS locales disponibles, consulte Regiones, Zonas de disponibilidad y Zonas locales.

  Actualmente, Zonas locales no es compatible con Cuadernos de Amazon EMR y no admite conexiones directas a Amazon EMR mediante el punto de conexión de VPC de la interfaz (AWS PrivateLink).

Cambios, mejoras y problemas resueltos

- Compatibilidad mejorada de aplicaciones para clústeres de alta disponibilidad
  - Para más información, consulte Aplicaciones admitidas en un clúster de EMR con varios nodos principales en la Guía de administración de Amazon EMR.
- Spark
  - Optimizaciones de rendimiento
- Hive
  - Optimizaciones de rendimiento
- Presto
  - Optimizaciones de rendimiento

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.27.0

Las siguientes notas de la versión incluyen información sobre la versión 5.27.0 de Amazon EMR. Los cambios son respecto a la versión 5.26.0.

Fecha de lanzamiento inicial: 23 de septiembre de 2019

Actualizaciones

- AWS SDK para Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- TensorFlow 1.14.0
- Conectores y controladores:
  - Conector de DynamoDB 4.12.0

Nuevas características

- (24 de octubre de 2019) Las siguientes características nuevas de Cuadernos de EMR están disponibles en todas las versiones de Amazon EMR.
  - Puede asociar repositorios de Git con EMR Notebooks para guardar los cuadernos en un entorno con control de versiones. Puede compartir código con sus compañeros y reutilizar los cuadernos de Jupyter existentes a través de repositorios de Git remotos. Para más información, consulte Agregar repositorios de Git con Cuadernos de Amazon EMR en la Guía de administración de Amazon EMR.
  - La utilidad nbdime ahora está disponible en Cuadernos de EMR para simplificar la comparación y la combinación de cuadernos.
  - Ahora son compatibles los cuadernos EMR. JupyterLab JupyterLab es un entorno de desarrollo interactivo basado en la web totalmente compatible con los portátiles Jupyter. Ahora puede elegir abrir su bloc de notas en el editor de cuadernos de JupyterLab Jupyter.
- (30 de octubre de 2019) Con las versiones 5.25.0 y posteriores de Amazon EMR, puede conectarse a la interfaz de usuario del servidor del historial de Spark desde la página Resumen del clúster o desde la pestaña Historial de aplicaciones de la consola. En lugar de configurar un proxy web a través de una conexión SSH, puede acceder rápidamente a la interfaz de usuario del servidor de historial de Spark para ver las métricas de las aplicaciones y acceder a los archivos de registro relevantes de los clústeres activos y terminados. Para más información, consulte Acceso fuera del clúster a las interfaces de usuario de aplicaciones persistentes en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Clúster de Amazon EMR con varios nodos principales
  - Puede instalar y ejecutar Flink en un clúster de Amazon EMR con varios nodos principales. Para más información, consulte Aplicaciones y características compatibles.
  - Puede configurar el cifrado transparente de HDFS en un clúster de Amazon EMR con varios nodos principales. Para más información, consulte Cifrado transparente de HDFS en clústeres de EMR con varios nodos principales.
  - Ahora puede modificar la configuración de las aplicaciones que se ejecutan en un clúster de Amazon EMR con varios nodos principales. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.
- Conector de DynamoDB para Amazon EMR
  - El conector de DynamoDB para Amazon EMR ahora admite los siguientes tipos de datos de DynamoDB: boolean, list, map, item y null. Para más información, consulte Configurar una tabla de Hive para ejecutar comandos de Hive.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.26.0

Las siguientes notas de la versión incluyen información sobre la versión 5.26.0 de Amazon EMR. Los cambios son respecto a la versión 5.25.0.

Fecha de lanzamiento inicial: 8 de agosto de 2019

Fecha de la última actualización: 19 de agosto de 2019

Actualizaciones

- AWS SDK para Java 1.11.595
- HBase 1.4,10
- Phoenix 4.14.2
- Conectores y controladores:
  - Conector de DynamoDB 4.11.0
  - MariaDB Connector 2.4.2
  - Controlador JDBC de Amazon Redshift, versión 1.2.32.1056

Nuevas características

- (Beta) Con Amazon EMR 5.26.0, puede lanzar un clúster que se integre con Lake Formation. Esta integración proporciona un acceso detallado a nivel de columnas a las bases de datos y tablas del catálogo de datos de Glue. AWS También permite un inicio de sesión único federado en cuadernos de EMR o Apache Zeppelin desde un sistema de identidades empresariales. Para obtener más información, consulte Integración de Amazon EMR con AWS Lake Formation (beta).

- (19 de agosto de 2019) El bloqueo de acceso público de Amazon EMR ya está disponible en todas las versiones de Amazon EMR que admiten grupos de seguridad. Bloquear el acceso público es una configuración que abarca toda la cuenta y que se aplica a cada región. AWS Bloquear el acceso público impide que se inicie un clúster cuando cualquier grupo de seguridad asociado al clúster tiene una regla que permite el tráfico entrante desde IPv4 0.0.0.0/0 o IPv6 : :/0 (acceso público) en un puerto, a menos que se especifique un puerto como excepción. El puerto 22 es una excepción de forma predeterminada. Para más información, consulte Uso del bloqueo de acceso público de Amazon EMR en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- EMR Notebooks
  - Con la versión 5.26.0 y posteriores de EMR, Cuadernos de EMR admite bibliotecas de Python limitadas a cuadernos, además de las bibliotecas de Python predeterminadas. Puede instalar bibliotecas limitadas a cuadernos desde el editor de cuadernos sin tener que volver a crear un clúster ni volver a asociar un cuaderno a un clúster. Las bibliotecas con ámbito de cuaderno se instalan en un entorno virtual de Python, por lo que solo se aplican a la sesión de cuaderno actual. Esto le permite aislar las dependencias de los cuadernos. Para más información, consulte Uso de bibliotecas adaptadas a los cuadernos en la Guía de administración de Amazon EMR.
- EMRFS
  - Puede habilitar una función de ETag verificación (beta) configurándola en. `fs.s3.consistent.metadata.etag.verification.enabled true` Con esta función, EMRFS utiliza Amazon S3 ETags para comprobar que los objetos que se están leyendo son de la última versión disponible. Esta función resulta útil para los casos de read-after-update uso en los que los archivos de Amazon S3 se sobrescriben y conservan el mismo nombre. Esta capacidad de ETag verificación no funciona actualmente con S3 Select. Para más información, consulte Configuración de la vista consistente.
- Spark
  - Las siguientes optimizaciones ahora se habilitan de forma predeterminada: la eliminación dinámica de particiones, DISTINCT antes de INTERSECT, las mejoras en la inferencia estadística del plan de SQL para JOIN seguida de consultas DISTINCT, el aplanamiento de las subconsultas escalares, el reordenamiento optimizado de las uniones y la unión mediante filtros de Bloom. Para más información, consulte Optimización del rendimiento de Spark.
  - Se ha mejorado la generación de código en todas las etapas para Sort Merge Join.

  - Se ha mejorado la reutilización de fragmentos de consultas y subconsultas.
  - Mejoras en la preasignación de los ejecutores en el inicio de Spark.
  - Las uniones con filtro de Bloom ya no se aplican cuando el lado más pequeño de la unión incluye una sugerencia de transmisión.
- Tez
  - Se ha resuelto un problema con Tez. La interfaz de usuario de Tez ahora funciona en un clúster de Amazon EMR con varios nodos principales.

Problemas conocidos

- Las capacidades mejoradas de generación de código en todas las etapas de Sort Merge Join pueden aumentar la presión de la memoria cuando están habilitadas. Esta optimización mejora el rendimiento, pero puede provocar reintentos de trabajo o errores si `spark.yarn.executor.memoryOverheadFactor` no se ajusta para proporcionar suficiente memoria. Para deshabilitar esta característica, establezca `spark.sql.sortMergeJoinExec.extendedCodegen.enabled` como “Falso”.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.25.0

Las siguientes notas de la versión incluyen información sobre la versión 5.25.0 de Amazon EMR. Los cambios son respecto a la versión 5.24.1.

Fecha de lanzamiento inicial: 17 de julio de 2019

Fecha de la última actualización: 30 de octubre de 2019

Amazon EMR 5.25.0

Actualizaciones

- AWS SDK para Java 1.11.566
- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- ZooKeeper 3.4.14

Nuevas características

- (30 de octubre de 2019) A partir de la versión 5.25.0 de Amazon EMR, podrá conectarse a la IU del servidor del historial de Spark desde la página Resumen del clúster o desde la pestaña Historial de aplicaciones de la consola. En lugar de configurar un proxy web a través de una conexión SSH, puede acceder rápidamente a la interfaz de usuario del servidor de historial de Spark para ver las métricas de las aplicaciones y acceder a los archivos de registro relevantes de

los clústeres activos y terminados. Para más información, consulte Acceso fuera del clúster a las interfaces de usuario de aplicaciones persistentes en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha mejorado el rendimiento de algunas uniones mediante el uso de filtros de Bloom para hacer un filtrado previo de las entradas. La optimización está deshabilitada de forma predeterminada y se puede habilitar configurando el parámetro `spark.sql.bloomFilterJoin.enabled` de configuración de Spark en `true`.
  - Se ha mejorado el rendimiento de la agrupación de columnas por tipo de cadena.
  - Se mejoró la configuración predeterminada de memoria y núcleos del ejecutor Spark de los tipos de instancias R4 para clústeres sin instalación. HBase
  - Se ha resuelto un problema anterior relacionado con la característica de eliminación dinámica de particiones, por el que la tabla eliminada tenía que estar en el lado izquierdo de la unión.
  - Se ha mejorado la optimización de DISTINCT antes de INTERSECT para aplicarla a casos adicionales relacionados con alias.
  - Se ha mejorado la inferencia estadística del plan de SQL para JOIN seguida de consultas DISTINCT. Esta mejora se ha desactivado de forma predeterminada y se puede habilitar configurando el parámetro de configuración de Spark `spark.sql.statsImprovements.enabled` como `true`. La característica Distinct antes de Intersect requiere esta optimización y se habilitará automáticamente cuando `spark.sql.optimizer.distinctBeforeIntersect.enabled` se establezca en `true`.
  - Se ha optimizado el orden de unión con base en el tamaño de la tabla y de los filtros. La optimización se ha desactivado de forma predeterminada y se puede habilitar configurando el parámetro de configuración de Spark `spark.sql.optimizer.sizeBasedJoinReorder.enabled` en `true`.

  Para más información, consulte Optimización del rendimiento de Spark.
- EMRFS
  - La configuración de EMRFS, `fs.s3.buckets.create.enabled`, ahora está deshabilitada de forma predeterminada. Tras realizar pruebas, se ha descubierto que al deshabilitar esta configuración se mejora el rendimiento y se evita la creación involuntaria de buckets de S3. Si su aplicación se basa en esta funcionalidad, puede habilitarla estableciendo la propiedad `fs.s3.buckets.create.enabled` en `true` en la clasificación de configuración `emrfs-`

`site`. Para obtener información, consulte Suministro de una configuración durante la creación de un clúster.

- Mejoras en el cifrado de disco local y en el cifrado de S3 en las configuraciones de seguridad (5 de agosto de 2019)
  - Se ha separado la configuración del cifrado de Amazon S3 de la configuración del cifrado de disco local en la configuración de seguridad.
  - Se ha agregado una opción para habilitar el cifrado de EBS en la versión 5.24.0 y posteriores. Al seleccionar esta opción, se cifra el volumen del dispositivo raíz además de los volúmenes de almacenamiento. Las versiones anteriores requerían el uso de una AMI personalizada para cifrar el volumen del dispositivo raíz.
  - Para más información, consulte Opciones de cifrado en la Guía de administración de Amazon EMR.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.24.1

Las siguientes notas de la versión incluyen información sobre la versión 5.24.1 de Amazon EMR. Los cambios son respecto a la versión 5.24.0.

Fecha de lanzamiento inicial: 26 de junio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.

- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de hadoop.

```
kinit -kt <keytab_file> <principal>
```

  Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.24.0

Las siguientes notas de la versión incluyen información sobre la versión 5.24.0 de Amazon EMR. Los cambios son respecto a la versión 5.23.0.

Fecha de lanzamiento inicial: 11 de junio de 2019

Fecha de la última actualización: 5 de agosto de 2019

Actualizaciones

- Flink 1.8.0
- Hue
- JupyterHub 0.9.6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK para Java 1.11,546
- Conectores y controladores:
  - Conector de DynamoDB 4.9.0

- MariaDB Connector 2.4.1
- Controlador JDBC de Amazon Redshift, versión 1.2.27.1051

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha agregado una optimización para eliminar las particiones de forma dinámica. La optimización se deshabilita de forma predeterminada. Para habilitarla, defina el parámetro de configuración de Spark `spark.sql.dynamicPartitionPruning.enabled` en `true`.
  - Se ha mejorado el rendimiento de las consultas `INTERSECT`. Esta optimización se deshabilita de forma predeterminada. Para habilitarla, defina el parámetro de configuración de Spark `spark.sql.optimizer.distinctBeforeIntersect.enabled` en `true`.
  - Se ha agregado una optimización para aplanar las subconsultas escalares con agregados que utilicen la misma relación. La optimización se deshabilita de forma predeterminada. Para habilitarla, defina el parámetro de configuración de Spark `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` en `true`.
  - Se ha mejorado la generación de código en todas las etapas.

  Para más información, consulte Optimización del rendimiento de Spark.
- Mejoras en el cifrado de disco local y en el cifrado de S3 en las configuraciones de seguridad (5 de agosto de 2019)
  - Se ha separado la configuración del cifrado de Amazon S3 de la configuración del cifrado de disco local en la configuración de seguridad.
  - Se ha agregado una opción para habilitar el cifrado de EBS. Al seleccionar esta opción, se cifra el volumen del dispositivo raíz además de los volúmenes de almacenamiento. Las versiones anteriores requerían el uso de una AMI personalizada para cifrar el volumen del dispositivo raíz.
  - Para más información, consulte Opciones de cifrado en la Guía de administración de Amazon EMR.

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.23.0

Las siguientes notas de la versión incluyen información sobre la versión 5.23.0 de Amazon EMR. Los cambios son respecto a la versión 5.22.0.

Fecha de lanzamiento inicial: 1 de abril de 2019

Fecha de la última actualización: 30 de abril de 2019

Actualizaciones

- AWS SDK para Java 1.11,519

Nuevas características

- (30 de abril de 2019) Con Amazon EMR 5.23.0 y versiones posteriores, puede lanzar un clúster con tres nodos principales para admitir la alta disponibilidad de aplicaciones como YARN Resource Manager, HDFS, Spark NameNode, Hive y Ganglia. El nodo principal ya no es un posible punto único de error con esta característica. Si se produce un error en uno de los nodos principales, se produce un error automáticamente en Amazon EMR a través de un nodo principal en espera y se sustituye el nodo principal con error por uno nuevo con las mismas acciones de arranque y configuración. Para más información, consulte Planificación y configuración de nodos principales.

Problemas conocidos

- Interfaz de usuario de Tez (corregida en la versión 5.26.0 de Amazon EMR)

  La IU de Tez no funciona en un clúster de EMR con varios nodos principales.
- Hue (corregido en la versión 5.24.0 de Amazon EMR)
  - Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Para evitar que aparezca el mensaje de error de Solr:

    1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
    2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

```
appblacklist = search
```

4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

```
sudo stop hue; sudo start hue
```

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

  > **Note**
  >
  > Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.22.0

Las siguientes notas de la versión incluyen información sobre la versión 5.22.0 de Amazon EMR. Los cambios son respecto a la versión 5.21.0.

> **Important**
>
> A partir de la versión 5.22.0 de Amazon EMR, Amazon EMR AWS utiliza Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Las versiones anteriores de Amazon EMR utilizan AWS Signature Version 2 en algunos casos, a menos que las notas de la versión indiquen que se utiliza exclusivamente Signature Version 4. Para obtener más información, consulte Autenticación de solicitudes (versión de AWS firma 4) y Autenticación de solicitudes (versión de AWS firma 2) en la Guía para desarrolladores de Amazon Simple Storage Service.

Fecha de lanzamiento inicial: 20 de marzo de 2019

Actualizaciones

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- Conectores y controladores:
  - Conector de DynamoDB 4.8.0
  - MariaDB Connector 2.2.6
  - Controlador JDBC de Amazon Redshift, versión 1.2.20.1043

Nuevas características

- Se modificó la configuración de EBS predeterminada para los tipos de EC2 instancias con almacenamiento exclusivo para EBS. Al crear un clúster utilizando la versión 5.22.0 y posteriores de Amazon EMR, la cantidad predeterminada de almacenamiento de EBS aumenta en función del tamaño de la instancia. Además, podemos dividir un mayor almacenamiento en varios

volúmenes, lo que ofrece un mayor desempeño de IOPS. Si desea utilizar una configuración de almacenamiento para instancias de EBS diferente, puede especificarlo al crear un clúster de EMR o añadir nodos a un clúster existente. Para más información sobre la cantidad de almacenamiento y la cantidad de volúmenes asignados de forma predeterminada a cada tipo de instancia, consulte Almacenamiento predeterminado de EBS para instancias en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha introducido una nueva propiedad de configuración para Spark en YARN, `spark.yarn.executor.memory0verheadFactor`. El valor de esta propiedad es un factor de escala que establece el valor de la sobrecarga de memoria en un porcentaje de la memoria del ejecutor, con un mínimo de 384 MB. Si la sobrecarga de memoria se establece de forma explícita mediante `spark.yarn.executor.memory0verhead`, esta propiedad no tendrá ningún efecto. El valor predeterminado es `0.1875`, que representa el 18,75 %. Este valor predeterminado para Amazon EMR deja más espacio en los contenedores de YARN para la sobrecarga de memoria del ejecutor que el 10 % predeterminado que ha establecido Spark internamente. El valor predeterminado de Amazon EMR del 18,75 % mostró empíricamente menos errores relacionados con la memoria en los puntos de referencia de TPC-DS.
  - Se ha agregado portabilidad con versiones anteriores para SPARK-26316 a fin de mejorar el rendimiento.
- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.

Problemas conocidos

- Hue (corregido en la versión 5.24.0 de Amazon EMR)
  - Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

```
Solr server could not be contacted properly:
HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
port=1978): Max retries exceeded with url: /solr/admin/info/
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Para evitar que aparezca el mensaje de error de Solr:

1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

   ```
   sudo vim /etc/hue/conf/hue.ini
   ```

3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

   ```
   appblacklist = search
   ```

4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

   ```
   sudo stop hue; sudo start hue
   ```

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.21.1

Las siguientes notas de la versión incluyen información sobre la versión 5.21.1 de Amazon EMR. Los cambios son respecto a la versión 5.21.0.

Fecha de lanzamiento inicial: 18 de julio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Problemas conocidos

- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

- SSH como usuario de hadoop con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de hadoop.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en /etc/hadoop.keytab y la entidad principal tiene la forma de hadoop/<hostname>@<REALM>.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.21.0

Las siguientes notas de la versión incluyen información sobre la versión 5.21.0 de Amazon EMR. Los cambios son respecto a la versión 5.20.0.

Fecha de lanzamiento inicial: 18 de febrero de 2019

Fecha de la última actualización: 3 de abril de 2019

Actualizaciones

- Flink 1.7.0
- Presto 0.215
- AWS SDK para Java 1.11.479

Nuevas características

- (3 de abril de 2019) Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS

Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Cambios, mejoras y problemas resueltos

- Zeppelin
  - Se ha adaptado ZEPPELIN-3878.

Problemas conocidos

- Hue (corregido en la versión 5.24.0 de Amazon EMR)
  - Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Para evitar que aparezca el mensaje de error de Solr:

    1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
    2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

       ```
       appblacklist = search
       ```

    4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

       ```
       sudo stop hue; sudo start hue
       ```

- Tez
  - Este problema se ha solucionado en la versión 5.22.0 de Amazon EMR.

Cuando se conecta a la interfaz de usuario de Tez en http: //:8080/tez-ui *MasterDNS* a través de una conexión SSH con el nodo principal del clúster, aparece el error «Falló la operación del adaptador; el servidor Timeline (ATS) está fuera de su alcance». Either it is down, or CORS is not enabled" o bien las tareas muestran inesperadamente N/A.

Esto se debe a que la interfaz de usuario de Tez realiza solicitudes al servidor YARN Timeline con `localhost` en vez de con el nombre de host del nodo principal. Como solución alternativa, hay un script disponible para ejecutar como acción o paso de arranque. El script actualiza el nombre del host en el archivo `configs.env` de Tez. Para más información y obtener la ubicación del script, consulte Bootstrap Instructions.

- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

  Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

  Solución:

  - SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
  - Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.20.0

Las siguientes notas de la versión incluyen información sobre la versión 5.20.0 de Amazon EMR. Los cambios son respecto a la versión 5.19.0.

Fecha de lanzamiento inicial: 18 de diciembre de 2018

Fecha de la última actualización: 22 de enero de 2019

Actualizaciones

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK para Java 1.11.461

Nuevas características

- (22 de enero de 2019) Se ha mejorado Kerberos en Amazon EMR para admitir la autenticación de las entidades principales desde un KDC externo. Esto centraliza la administración de entidades principales porque varios clústeres pueden compartir un único KDC externo. Además, el KDC externo puede tener una relación de confianza entre ámbitos con un dominio de Active Directory.

Esto permite que todos los clústeres autentiquen entidades principales desde Active Directory. Para más información, consulte Uso de la autenticación de Kerberos en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- AMI de Amazon Linux predeterminada para Amazon EMR
  - Se ha actualizado el paquete de Python3 de Python 3.4 a 3.6.
- El confirmador optimizado para S3 de EMRFS
  - El confirmador optimizado para S3 de EMRFS ahora está habilitado de forma predeterminada, lo que mejora el rendimiento de escritura. Para obtener más información, consulte Uso del confirmador optimizado para S3 de EMRFS.
- Hive
  - Se ha adaptado HIVE-16686.
- Glue con Spark y Hive
  - En EMR 5.20.0 o versiones posteriores, la reducción de particiones en paralelo se habilita automáticamente para Spark y Hive cuando se utiliza AWS Glue Data Catalog como metaalmacén. Este cambio reduce significativamente el tiempo de planificación de consultas al ejecutar varias solicitudes en paralelo para recuperar particiones. El número total de segmentos que se pueden ejecutar simultáneamente oscila entre 1 y 10. El valor predeterminado es 5, que es la configuración recomendada. Para cambiarlo, especifique la propiedad `aws.glue.partition.num.segments` en la clasificación de la configuración `hive-site`. Si se produce una limitación, puede desactivar la característica cambiando el valor a 1. Para más información, consulte Estructura Segment de AWS Glue.

Problemas conocidos

- Hue (corregido en la versión 5.24.0 de Amazon EMR)
  - Hue en ejecución en Amazon EMR no es compatible con Solr. A partir de la versión 5.20.0 de Amazon EMR, un problema de configuración incorrecta hace que Solr se habilite y aparezca un mensaje de error inofensivo similar al siguiente:

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    ```

```
system?user.name=hue&doAs=administrator&wt=json (Caused by
NewConnectionError(': Failed to establish a new connection: [Errno 111]
Connection refused',))
```

Para evitar que aparezca el mensaje de error de Solr:

1. Conéctese a la línea de comandos del nodo principal utilizando SSH.
2. Use un editor de texto para abrir el archivo `hue.ini`. Por ejemplo:

   ```
   sudo vim /etc/hue/conf/hue.ini
   ```

3. Busque el término `appblacklist` y modifique la línea de la siguiente manera:

   ```
   appblacklist = search
   ```

4. Guarde los cambios y reinicie Hue como se muestra en el siguiente ejemplo:

   ```
   sudo stop hue; sudo start hue
   ```

- Tez
  - Este problema se ha solucionado en la versión 5.22.0 de Amazon EMR.

    Cuando te conectas a la interfaz de usuario de Tez en http: //:8080/tez-ui *`MasterDNS`* mediante una conexión SSH con el nodo principal del clúster, aparece el error «Falló la operación del adaptador; el servidor Timeline (ATS) está fuera de alcance». Either it is down, or CORS is not enabled” o bien las tareas muestran inesperadamente N/A.

    Esto se debe a que la interfaz de usuario de Tez realiza solicitudes al servidor YARN Timeline con `localhost` en vez de con el nombre de host del nodo principal. Como solución alternativa, hay un script disponible para ejecutar como acción o paso de arranque. El script actualiza el nombre del host en el archivo `configs.env` de Tez. Para más información y obtener la ubicación del script, consulte Bootstrap Instructions.
- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.
- Problema conocido en clústeres con varios nodos principales y autenticación de Kerberos

Si ejecuta clústeres con varios nodos principales y autenticación de Kerberos en las versiones 5.20.0 y posteriores de Amazon EMR, es posible que tenga problemas con las operaciones del clúster, como la reducción vertical o el envío escalonado, después de que el clúster se haya estado ejecutando durante algún tiempo. El periodo de tiempo depende del periodo de validez del ticket de Kerberos que se haya definido. El problema de la reducción vertical afecta tanto a las solicitudes de reducción vertical automática como a las solicitudes de reducción vertical explícita que haya enviado. Las operaciones adicionales del clúster también pueden verse afectadas.

Solución:

- SSH como usuario de `hadoop` con el nodo principal líder del clúster de EMR con varios nodos principales.
- Ejecute el siguiente comando para renovar el ticket de Kerberos para el usuario de `hadoop`.

```
kinit -kt <keytab_file> <principal>
```

Normalmente, el archivo keytab se encuentra en `/etc/hadoop.keytab` y la entidad principal tiene la forma de `hadoop/<hostname>@<REALM>`.

> **Note**
>
> Esta solución alternativa estará en vigor durante el periodo de validez del ticket de Kerberos. Esta duración es de 10 horas de forma predeterminada, pero se puede configurar con los ajustes de Kerberos. Debe volver a ejecutar el comando anterior una vez que venza el ticket de Kerberos.

Versión 5.19.0

Las siguientes notas de la versión incluyen información sobre la versión 5.19.0 de Amazon EMR. Los cambios son respecto a la versión 5.18.0.

Fecha de lanzamiento inicial: 7 de noviembre de 2018

Fecha de la última actualización: 19 de noviembre de 2018

Actualizaciones

- Hadoop 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- ZooKeeper 3.4.13
- AWS SDK para Java 1.11.433

Nuevas características

- (19 de noviembre de 2018) Cuadernos de EMR es un entorno administrado basado en cuadernos de Jupyter. Es compatible con los núcleos mágicos de Spark para PySpark Spark SQL, Spark R y Scala. Cuadernos de EMR se puede utilizar con clústeres creados con la versión 5.18.0 y posteriores de Amazon EMR. Para más información, consulte Uso de Cuadernos de EMR en la Guía de administración de Amazon EMR.
- El confirmador optimizado para S3 de EMRFS está disponible al escribir archivos Parquet con Spark y EMRFS. Este confirmador mejora el rendimiento de escritura. Para obtener más información, consulte Uso del confirmador optimizado para S3 de EMRFS.

Cambios, mejoras y problemas resueltos

- YARN
  - Se ha modificado la lógica que limita el proceso maestro de la aplicación a ejecutarse en los nodos principales. Esta funcionalidad ahora utiliza la característica y las propiedades de las etiquetas de nodos de YARN en las clasificaciones de configuración `yarn-site` y `capacity-scheduler`. Para obtener información, consulte https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN.
- AMI de Amazon Linux predeterminada para Amazon EMR
  - `ruby18`, `php56` y `gcc48` ya no se instalan de forma predeterminada. Se pueden instalar si se desea utilizando `yum`.

- La gema aws-sdk ruby ya no se instala de forma predeterminada. Si lo desea, se puede instalar utilizando `gem install aws-sdk`. También se pueden instalar componentes específicos. Por ejemplo, `gem install aws-sdk-s3`.

Problemas conocidos

- Cuadernos de EMR: en algunas circunstancias, con varios editores de cuadernos abiertos, puede parecer que el editor de cuadernos no puede conectarse al clúster. Si esto ocurre, borre las cookies del navegador y, a continuación, vuelva a abrir los editores de cuadernos.
- CloudWatch ContainerPending Escalado métrico y automático: (corregido en la versión 5.20.0) Amazon EMR puede emitir un valor negativo para. `ContainerPending` Si `ContainerPending` se utiliza en una regla de escalado automático, el escalado automático no se comportará como se esperaba. Evite usar `ContainerPending` con el escalado automático.
- En las versiones 5.19.0, 5.20.0 y 5.21.0 de Amazon EMR, las etiquetas de los nodos de YARN se almacenan en un directorio de HDFS. En algunas situaciones, esto provoca demoras en el inicio del nodo principal y, a continuación, provoca que el clúster agote el tiempo de espera y se produzca un error en el lanzamiento. A partir de la versión 5.22.0 de Amazon EMR, este problema se ha resuelto. Las etiquetas de los nodos de YARN se almacenan en el disco local de cada nodo del clúster, lo que evita las dependencias de HDFS.

Versión 5.18.0

Las siguientes notas de la versión incluyen información sobre la versión 5.18.0 de Amazon EMR. Los cambios son respecto a la versión 5.17.0.

Fecha de lanzamiento inicial: 24 de octubre de 2018

Actualizaciones

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

Nuevas características

- A partir de Amazon EMR 5.18.0, puede utilizar el repositorio de artefactos de Amazon EMR para compilar el código de trabajos con las versiones exactas de las bibliotecas y dependencias que están disponibles con determinadas versiones de Amazon EMR. Para obtener más información, consulte Comprobación de dependencias mediante el repositorio de artefactos de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Hive
  - Se ha agregado soporte para S3 Select. Para obtener más información, consulte Uso de S3 Select con Hive para mejorar el rendimiento.
- Presto
  - Se ha agregado soporte para S3 Select Pushdown. Para obtener más información, consulte Uso de S3 Select Pushdown con Presto para mejorar el rendimiento.
- Spark
  - Se ha cambiado la configuración predeterminada de log4j para Spark y ahora permite distribuir los registros de contenedores cada hora para los trabajos de streaming de Spark. Esto ayuda a evitar que se eliminen los registros de los trabajos de streaming de Spark de ejecución prolongada.

Versión 5.17.1

Las siguientes notas de la versión incluyen información sobre la versión 5.17.1 de Amazon EMR. Los cambios son respecto a la versión 5.17.0.

Fecha de lanzamiento inicial: 18 de julio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Versión 5.17.0

Las siguientes notas de la versión incluyen información sobre la versión 5.17.0 de Amazon EMR. Los cambios son respecto a la versión 5.16.0.

Fecha de lanzamiento inicial: 30 de agosto de 2018

Actualizaciones

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

Nuevas características

- Se ha agregado compatibilidad con TensorFlow. Para obtener más información, consulte TensorFlow.

Cambios, mejoras y problemas resueltos

- JupyterHub
  - Se ha agregado compatibilidad con la persistencia de cuadernos en Amazon S3. Para obtener más información, consulte Configuración de la persistencia de los cuadernos en Amazon S3.
- Spark
  - Se ha agregado soporte para S3 Select. Para obtener más información, consulte Uso de S3 Select con Spark para mejorar el rendimiento de las consultas.
- Se han resuelto los problemas con las métricas de CloudWatch y la característica de escalado automático en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR.

Problemas conocidos

- Al crear un clúster que utiliza Kerberos con Livy instalado, Livy produce un error que indica que la autenticación simple no está habilitada. El problema se resuelve al reiniciar el servidor de Livy. Como solución alternativa, agregue un paso durante la creación del clúster que ejecuta `sudo restart livy-server` en el nodo principal.

- Si utiliza una AMI personalizada de Amazon Linux basada en una AMI de Amazon Linux con una fecha de creación del 11 de agosto de 2018, el servidor de Oozie no podrá iniciarse. Si utiliza Oozie, cree una AMI personalizada basada en un ID de AMI de Amazon Linux con una fecha de creación diferente. Puede utilizar el siguiente AWS CLI comando para obtener una lista de imágenes IDs para todas las HVM de Amazon Linux AMIs con una versión 2018.03, junto con la fecha de lanzamiento, de modo que pueda elegir una AMI de Amazon Linux adecuada como base. MyRegion Sustitúyalo por tu identificador de región, como us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

Versión 5.16.0

Las siguientes notas de la versión incluyen información sobre la versión 5.16.0 de Amazon EMR. Los cambios son respecto a la versión 5.15.0.

Fecha de lanzamiento inicial: 19 de julio de 2018

Actualizaciones

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK para Java 1.11.336
- CUDA 9.2
- Controlador JDBC de Amazon Redshift, versión 1.2 15.1025

Cambios, mejoras y problemas resueltos

- HBase

  - Se ha adaptado HBASE-20723
- Presto
  - Cambios de configuración para admitir la autenticación LDAP. Para obtener más información, consulte Uso de la autenticación LDAP para Presto en Amazon EMR.
- Spark
  - La versión 2.3.1 de Apache Spark, disponible a partir de la versión 5.16.0 de Amazon EMR, aborda CVE-2018-8024 y CVE-2018-1334. Se recomienda migrar las versiones anteriores de Spark a la versión 2.3.1 o posteriores de Spark.

Problemas conocidos

- Esta versión de lanzamiento no es compatible con los tipos de instancias c1.medium ni m1.small. Los clústeres que utilizan cualquiera de estos tipos de instancias no se inician. Como solución alternativa, especifique un tipo de instancia diferente o utilice una versión de lanzamiento diferente.
- Al crear un clúster que utiliza Kerberos con Livy instalado, Livy produce un error que indica que la autenticación simple no está habilitada. El problema se resuelve al reiniciar el servidor de Livy. Como solución alternativa, agregue un paso durante la creación del clúster que ejecuta `sudo restart livy-server` en el nodo principal.
- Una vez que se reinicie el nodo principal o se reinicie el controlador de instancias, no se recopilarán las CloudWatch métricas y la función de escalado automático no estará disponible en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR. Este problema se ha corregido en la versión 5.17.0 de Amazon EMR.

Versión 5.15.0

Las siguientes notas de la versión incluyen información sobre la versión 5.15.0 de Amazon EMR. Los cambios son respecto a la versión 5.14.0.

Fecha de lanzamiento inicial: 21 de junio de 2018

Actualizaciones

- Se actualizó a la versión 1.4.4 HBase
- Se ha actualizado Hive a la versión 2.3.3
- Se ha actualizado Hue a la versión 4.2.0

- Se ha actualizado Oozie a la versión 5.0.0
- Se ha actualizado ZooKeeper a la versión 3.4.12
- Se actualizó el AWS SDK a la versión 1.11.333

Cambios, mejoras y problemas resueltos

- Hive
  - Se ha adaptado HIVE-18069
- Hue
  - Se ha actualizado Hue para autenticarse correctamente con Livy cuando Kerberos está habilitado. Ahora se admite Livy cuando se usa Kerberos con Amazon EMR.
- JupyterHub
  - Se actualizó JupyterHub para que Amazon EMR instale las bibliotecas cliente LDAP de forma predeterminada.
  - Se ha corregido un error en el script que genera los certificados autofirmados.

Problemas conocidos

- Esta versión de lanzamiento no es compatible con los tipos de instancias c1.medium ni m1.small. Los clústeres que utilizan cualquiera de estos tipos de instancias no se inician. Como solución alternativa, especifique un tipo de instancia diferente o utilice una versión de lanzamiento diferente.
- Una vez que se reinicie el nodo principal o se reinicie el controlador de instancias, no se recopilarán las CloudWatch métricas y la función de escalado automático no estará disponible en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR. Este problema se ha corregido en la versión 5.17.0 de Amazon EMR.

Versión 5.14.1

Las siguientes notas de la versión incluyen información sobre la versión 5.14.1 de Amazon EMR. Los cambios son respecto a la versión 5.14.0.

Fecha de lanzamiento inicial: 17 de octubre de 2018

Se ha actualizado la AMI predeterminada de Amazon EMR para abordar posibles vulnerabilidades de seguridad.

Versión 5.14.0

Las siguientes notas de la versión incluyen información sobre la versión 5.14.0 de Amazon EMR. Los cambios son respecto a la versión 5.13.0.

Fecha de lanzamiento inicial: 4 de junio de 2018

Actualizaciones

- Se ha actualizado Flink a la versión 1.4.2
- Se actualizó Apache a la versión 1.1.0 MXnet
- Se ha actualizado Apache Sqoop a la versión 1.4.7

Nuevas características

- Se agregó JupyterHub soporte. Para obtener más información, consulte JupyterHub.

Cambios, mejoras y problemas resueltos

- EMRFS
  - La cadena userAgent de las solicitudes para Amazon S3 se ha actualizado para incluir la información de usuario y de grupo de la entidad principal que la invoca. Esto se puede usar con AWS CloudTrail los registros para un seguimiento más completo de las solicitudes.
- HBase
  - Se ha incluido HBASE-20447, que soluciona un problema que podía provocar problemas de caché, especialmente cuando las regiones estaban divididas.
- MXnet
  - Se han agregado bibliotecas OpenCV.
- Spark
  - Cuando Spark escribe archivos Parquet en una ubicación de Amazon S3 mediante EMRFS, el FileOutputCommitter algoritmo se ha actualizado para usar la versión 2 en lugar de la versión 1. Esto reduce el número de cambios de nombre, lo que mejora el rendimiento de la aplicación. Este cambio no afecta a:
    - Aplicaciones distintas de Spark.
    - Aplicaciones que escriben en otros sistemas de archivos, como HDFS (que aún utilizan la versión 1 de). FileOutputCommitter

- Aplicaciones que utilizan otros formatos de salida, como texto o csv, que ya utilizan la escritura directa de EMRFS.

Problemas conocidos

- JupyterHub
  - No se admite el uso de clasificaciones de configuración para configurar JupyterHub cuadernos de Jupyter individuales al crear un clúster. Edite manualmente los archivos jupyterhub_config.py y jupyter_notebook_config.py para cada usuario. Para obtener más información, consulte Configurando JupyterHub.
  - JupyterHub no se inicia en los clústeres de una subred privada y se produce un error con el mensaje. `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` Esto es provocado por un error en el script que genera los certificados autofirmados. Utilice la siguiente solución alternativa para generar certificados autofirmados. Todos los comandos se ejecutan mientras está conectado al nodo principal.

    1. Copie el script de generación de certificados del contenedor al nodo principal:

       ```
       sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
       ```

    2. Utilice un editor de texto para cambiar la línea 23 y cambiar el nombre de host público por un nombre de host local, como se muestra a continuación:

       ```
       local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
       ```

    3. Ejecute el script para generar certificados autofirmados:

       ```
       sudo bash ./gen_self_signed_cert.sh
       ```

    4. Mueva los archivos de certificados que genera el script al directorio `/etc/jupyter/conf/`:

       ```
       sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
       ```

    Puedes revisar `tail` el `jupyter.log` archivo para comprobar que se ha JupyterHub reiniciado y que devuelve un código de respuesta de 200. Por ejemplo:

    ```
    tail -f /var/log/jupyter/jupyter.log
    ```

Este debe devolver una respuesta similar a la siguiente:

```
# [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
 https://:9443/
# 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
```

- Una vez que se reinicie el nodo principal o se reinicie el controlador de instancias, no se recopilarán las CloudWatch métricas y la función de escalado automático no estará disponible en las versiones 5.14.0, 5.15.0 o 5.16.0 de Amazon EMR. Este problema se ha corregido en la versión 5.17.0 de Amazon EMR.

Versión 5.13.0

Las siguientes notas de la versión incluyen información sobre la versión 5.13.0 de Amazon EMR. Los cambios son respecto a la versión 5.12.0.

Actualizaciones

- Se ha actualizado Spark a la versión 2.3.0
- Se actualizó a la versión 1.4.2 HBase
- Se ha actualizado Presto a la versión 0.194
- Actualizado a la versión AWS SDK para Java 1.11.297

Cambios, mejoras y problemas resueltos

- Hive
  - Se ha adaptado HIVE-15436. Se ha mejorado la colmena APIs para que solo devuelva las vistas.

Problemas conocidos

- MXNet actualmente no tiene bibliotecas OpenCV.

Versión 5.12.2

Las siguientes notas de la versión incluyen información sobre la versión 5.12.2 de Amazon EMR. Los cambios son respecto a la versión 5.12.1.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- Esta versión aborda una posible vulnerabilidad de seguridad.

Versión 5.12.1

Las siguientes notas de la versión incluyen información sobre la versión 5.12.1 de Amazon EMR. Los cambios son respecto a la versión 5.12.0.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versión 5.12.0

Las siguientes notas de la versión incluyen información sobre la versión 5.12.0 de Amazon EMR. Los cambios son respecto a la versión 5.11.1.

Actualizaciones

- AWS SDK para Java 1.11.238 ⇒ 1.11.267. Para obtener más información, consulta el registro de cambios de AWS SDK for Java GitHub.
- Hadoop 2.7.3 ⇒ 2.8.3. Para más información, consulte Versiones de Apache Hadoop.
- Flink 1.3.2 ⇒ 1.4.0. Para más información, consulte Apache Flink 1.4.0 Release Announcement.
- HBase 1.3.1 ⇒ 1.4.0. Para obtener más información, consulta el anuncio de HBase lanzamiento.
- Hue 4.0.1 ⇒ 4.1.0. Para más información, consulte las notas de la versión.
- MxNet 0.12.0 ⇒ 1.0.0. Para obtener más información, consulte el registro de MXNet cambios. GitHub

- Presto 0.187 ⇒ 0.188. Para más información, consulte las notas de la versión.

Cambios, mejoras y problemas resueltos

- Hadoop
  - La propiedad `yarn.resourcemanager.decommissioning.timeout` se ha cambiado a `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Puede utilizar esta propiedad para personalizar la reducción vertical del clúster. Para más información, consulte Reducción vertical de clúster en la Guía de administración de Amazon EMR.
  - La CLI de Hadoop ha agregado la opción `-d` al comando `cp` (copia), que especifica la copia directa. Puede utilizarlo para evitar la creación de un archivo `.COPYING` intermediario, lo que agiliza la copia de datos entre Amazon S3. Para más información, consulte HADOOP-12384.
- Pig
  - Se ha agregado la clasificación de configuración `pig-env`, que simplifica la configuración de las propiedades del entorno de Pig. Para obtener más información, consulte Configuración de aplicaciones.
- Presto
  - Se ha agregado la clasificación de configuración `presto-connector-redshift`, que se puede utilizar para configurar valores en el archivo de configuración `redshift.properties` de Presto. Para más información, consulte Redshift Connector en la documentación de Presto y Configuración de aplicaciones.
  - Se ha agregado compatibilidad con Presto para EMRFS y es la configuración predeterminada. Las versiones anteriores de Amazon EMR utilizaban PrestoS3FileSystem, que era la única opción. Para obtener más información, consulte Configuración de EMRFS y PrestoS3PrestoS3 FileSystem.

    > **Note**
    >
    > Si consulta los datos subyacentes en Amazon S3 con la versión 5.12.0 de Amazon EMR, pueden producirse errores en Presto. Esto se debe a que Presto no recoge los valores de clasificación de configuración de `emrfs-site.xml`. Como solución alternativa, cree un subdirectorio de `emrfs` en `usr/lib/presto/plugin/hive-hadoop2/` y cree un enlace simbólico en `usr/lib/presto/plugin/hive-hadoop2/emrfs` al archivo `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existente. A continuación,

reinicie el proceso presto-server (`sudo presto-server stop` seguido de `sudo presto-server start`).

- Spark
  - SPARK-22036 compatible: la multiplicación a veces devuelve un valor nulo. BigDecimal

Problemas conocidos

- MXNet no incluye bibliotecas OpenCV.
- SparkR no está disponible para los clústeres creados con una AMI personalizada porque R no está instalado de forma predeterminada en los nodos del clúster.

Versión 5.11.3

Las siguientes notas de la versión incluyen información sobre la versión 5.11.3 de Amazon EMR. Los cambios son respecto a la versión 5.11.2.

Fecha de lanzamiento inicial: 18 de julio de 2019

Cambios, mejoras y problemas resueltos

- Se ha actualizado la AMI de Amazon Linux predeterminada para Amazon EMR a fin de incluir importantes actualizaciones de seguridad del kernel de Linux, incluido el problema de denegación de servicio TCP SACK (AWS-2019-005).

Versión 5.11.2

Las siguientes notas de la versión incluyen información sobre la versión 5.11.2 de Amazon EMR. Los cambios son respecto a la versión 5.11.1.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- Esta versión aborda una posible vulnerabilidad de seguridad.

Versión 5.11.1

Las siguientes notas de la versión incluyen información sobre la versión 5.11.1 de Amazon EMR. Los cambios son respecto a la versión 5.11.0 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Problemas conocidos

- MXNet no incluye bibliotecas OpenCV.
- Hive 2.3.2 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

  Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

Versión 5.11.0

Las siguientes notas de la versión incluyen información sobre la versión 5.11.0 de Amazon EMR. Los cambios son respecto a la versión 5.10.0 de Amazon EMR.

Actualizaciones

Las siguientes aplicaciones y componentes se han actualizado en esta versión para incluir las siguientes versiones.

- Hive 2.3.2
- Spark 2.2.1
- SDK para Java 1.11.238

Nuevas características

- Spark
  - Se ha añadido el ajuste `spark.decommissioning.timeout.threshold`, que mejora el comportamiento de retirada de Spark cuando se utilizan instancias de spot. Para obtener más información, consulte Configuración del comportamiento de retirada de nodos.
  - Se agregó el `aws-sagemaker-spark-sdk` componente a Spark, que instala Amazon SageMaker Spark y las dependencias asociadas para la integración de Spark con Amazon. SageMaker Puedes usar Amazon SageMaker Spark para crear canalizaciones de aprendizaje automático (ML) de Spark mediante Amazon SageMaker Stages. Para obtener más información, consulta el léame de SageMaker Spark sobre el GitHub uso de Apache Spark con Amazon SageMaker en la Guía para SageMaker desarrolladores de Amazon.

Problemas conocidos

- MXNet no incluye bibliotecas OpenCV.
- Hive 2.3.2 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

  Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

Versión 5.10.0

Las siguientes notas de la versión incluyen información sobre la versión 5.10.0 de Amazon EMR. Los cambios son respecto a la versión 5.9.0 de Amazon EMR.

Actualizaciones

Las siguientes aplicaciones y componentes se han actualizado en esta versión para incluir las siguientes versiones.

- AWS SDK para Java 1.11.21
- Hive 2.3.1
- Presto 0.187

Nuevas características

- Se ha agregado compatibilidad para autenticación Kerberos. Para más información, consulte Uso de la autenticación de Kerberos en la Guía de administración de Amazon EMR
- Se ha agregado compatibilidad con los roles de IAM de las solicitudes de EMRFS para Amazon S3. Para más información, consulte Configurar roles de IAM de solicitudes de EMRFS para Amazon S3 en la Guía de administración de Amazon EMR.
- Se ha agregado un soporte para los tipos de instancias P2 y P3 basados en GPU. Para obtener más información, consulte Instancias Amazon EC2 P2 e instancias Amazon EC2 P3. El controlador NVIDIA 384,81 y el controlador CUDA 9.0.176 están instalados en estos tipos de instancias de forma predeterminada.
- Se agregó compatibilidad con Apache MXNet.

Cambios, mejoras y problemas resueltos

- Presto
  - Se ha añadido compatibilidad con el uso del catálogo de datos de AWS Glue como metastore predeterminado de Hive. Para obtener más información, consulte Uso de Presto con el catálogo de datos de AWS Glue.
  - Se ha agregado compatibilidad para las funciones geoespaciales.
  - Se ha agregado compatibilidad con las uniones de derrame en el disco.
  - Se ha agregado compatibilidad con el conector de Redshift.
- Spark
  - Se ha adaptado SPARK-20640, lo que hace que el tiempo de espera de rpc y los reintentos de los valores de registro de reorganización sean configurables

mediante las propiedades `spark.shuffle.registration.timeout` y `spark.shuffle.registration.maxAttempts`.

- El SPARK-21549 es compatible, lo que corrige un error que se produce al grabar archivos personalizados en ubicaciones que no son HDFS. OutputFormat

- Se ha adaptado Hadoop-13270
- Se han eliminado las bibliotecas Numpy, Scipy y Matplotlib de la AMI base de Amazon EMR. Si la aplicación requiere estas bibliotecas, están disponibles en el repositorio de aplicaciones, por lo que puede utilizar una acción de arranque para instalarlas en todos los nodos mediante `yum install`.
- La AMI base de Amazon EMR ya no incluye paquetes RPM de aplicaciones, por lo que los paquetes RPM ya no están presentes en los nodos del clúster. Custom AMIs y la AMI base de Amazon EMR ahora hacen referencia al repositorio de paquetes RPM de Amazon S3.
- Gracias a la introducción de la facturación por segundo en Amazon EC2, el comportamiento de reducción predeterminado ahora es Finalizar al finalizar la tarea en lugar de Finalizar a la hora de la instancia. Para más información, consulte Reducción de escala del clúster.

Problemas conocidos

- MXNet no incluye bibliotecas OpenCV.
- Hive 2.3.1 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

  Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

Versión 5.9.0

Las siguientes notas de la versión incluyen información sobre la versión 5.9.0 de Amazon EMR. Los cambios son respecto a la versión 5.8.0 de Amazon EMR.

Fecha de lanzamiento: 5 de octubre de 2017

Última actualización de características: 12 de octubre de 2017

Actualizaciones

Las siguientes aplicaciones y componentes se han actualizado en esta versión para incluir las siguientes versiones.

- AWS SDK para Java versión 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

Nuevas características

- Se ha añadido compatibilidad con Livy (versión 0.4.0-incubating). Para obtener más información, consulte Apache Livy.
- Se ha añadido compatibilidad con Hue Notebook para Spark.
- Se agregó soporte para las EC2 instancias Amazon de la serie i3 (12 de octubre de 2017).

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha añadido un nuevo conjunto de características que ayudan a asegurarse de que Spark gestiona de un modo más correcto la terminación de los nodos debido a una solicitud de cambio de tamaño manual o a una política de escalado automático. Para obtener más información, consulte Configuración del comportamiento de retirada de nodos.
  - Se utiliza SSL en lugar de 3DES para el cifrado en tránsito del servicio de transferencia de bloques, lo que mejora el rendimiento cuando se utilizan tipos de EC2 instancias de Amazon con AES-NI.
  - Se ha adaptado SPARK-21494.
- Zeppelin
  - Se ha adaptado ZEPPELIN-2377.
- HBase

  - Se agregó el parche HBASE-18533, que permite valores adicionales para la configuración mediante la clasificación de configuración. HBase BucketCache `hbase-site`
- Hue
  - Se agregó compatibilidad con AWS Glue Data Catalog para el editor de consultas Hive en Hue.
  - Los superusuarios de Hue pueden acceder, de forma predeterminada, a todos los archivos para los que tienen autorización de acceso los roles de IAM de Amazon EMR. Los usuarios recién creados no obtienen automáticamente permisos de acceso al navegador de archivos de Amazon S3, por lo que es necesario activar los permisos `filebrowser.s3_access` para su grupo.
- Se ha solucionado un problema que impedía el acceso a los datos subyacentes de JSON creados con el Catálogo de datos de AWS Glue.

Problemas conocidos

- El lanzamiento del clúster produce un error cuando se instalan todas las aplicaciones y no se cambia el tamaño del volumen raíz de Amazon EBS predeterminado. Como solución alternativa, utilice el `aws emr create-cluster` comando del AWS CLI y especifique un parámetro más grande`--ebs-root-volume-size`.
- Hive 2.3.0 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

  Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.

Versión 5.8.2

Las siguientes notas de la versión incluyen información sobre la versión 5.8.2 de Amazon EMR. Los cambios son respecto a la versión 5.8.1.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versión 5.8.1

Las siguientes notas de la versión incluyen información sobre la versión 5.8.1 de Amazon EMR. Los cambios son respecto a la versión 5.8.0 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versión 5.8.0

Las siguientes notas de la versión incluyen información sobre la versión 5.8.0 de Amazon EMR. Los cambios son respecto a la versión 5.7.0 de Amazon EMR.

Fecha de lanzamiento inicial: 10 de agosto de 2017

Última actualización de características: 25 de septiembre de 2017

Actualizaciones

Las siguientes aplicaciones y componentes se han actualizado en esta versión para incluir las siguientes versiones:

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0. Para más información, consulte Notas de la versión en el sitio de Apache Hive.
- Spark 2.2.0. Para más información, consulte Notas de la versión en el sitio de Apache Spark.

Nuevas características

- Se ha agregado la posibilidad de visualizar el historial de aplicaciones (25 de septiembre de 2017). Para más información, consulte Visualización del historial de aplicaciones en la Guía de administración de Amazon EMR.

Cambios, mejoras y problemas resueltos

- Integración con AWS Glue Data Catalog
  - Se ha añadido la posibilidad de que Hive y Spark SQL utilicen AWS Glue Data Catalog como almacén de metadatos de Hive. Para obtener más información, consulte Se ha agregado compatibilidad para utilizar el Catálogo de datos de AWS Glue como metaalmacén predeterminado de Hive y Usa el catálogo de AWS Glue Data Catalog con Spark en Amazon EMR.
- Se ha añadido el Application history (Historial de aplicaciones) a los detalles del clúster, lo que le permite ver datos históricos de las aplicaciones de YARN y detalles adicionales de las aplicaciones de Spark. Para obtener más información, consulte Ver el historial de aplicaciones en la Guía de administración de Amazon EMR.
- Oozie
  - Se ha adaptado OOZIE-2748.
- Hue
  - Se ha adaptado HUE-5859
- HBase
  - Se agregó un parche para mostrar la hora de inicio del servidor HBase maestro mediante el uso de Java Management Extensions (JMX). `getMasterInitializedTime`
  - Se ha añadido un parche que mejora la hora de inicio del clúster.

Problemas conocidos

- El lanzamiento del clúster produce un error cuando se instalan todas las aplicaciones y no se cambia el tamaño del volumen raíz de Amazon EBS predeterminado. Como solución alternativa, utilice el `aws emr create-cluster` comando del AWS CLI y especifique un parámetro mayor`--ebs-root-volume-size`.

- Hive 2.3.0 establece `hive.compute.query.using.stats=true` de forma predeterminada. Esto desemboca en consultas que obtienen datos de las estadísticas existentes en lugar de directamente de los datos, lo que puede dar lugar a confusión. Por ejemplo, si tiene una tabla con `hive.compute.query.using.stats=true` y carga nuevos archivos en la tabla `LOCATION`, la ejecución de una consulta `SELECT COUNT(*)` en la tabla devuelve el recuento de las estadísticas, en lugar de seleccionar las filas añadidas.

  Para resolver este problema, utilice el comando `ANALYZE TABLE` para recopilar nuevas estadísticas o establezca `hive.compute.query.using.stats=false`. Para obtener más información, consulte Statistics in Hive en la documentación de Apache Hive.
- Spark: cuando se utiliza Spark, hay un problema de fuga de un controlador de archivos con el daemon apppusher que puede aparecer con un trabajo de Spark de ejecución prolongada después de varias horas o días. Para solucionar el problema, conéctese al nodo principal y escriba `sudo /etc/init.d/apppusher stop`. Esto detiene el daemon apppusher, que Amazon EMR reiniciará automáticamente.
- Historial de aplicaciones
  - Los datos históricos de los ejecutores de Spark inactivos no están disponibles.
  - El historial de aplicaciones no está disponible para clústeres que utilizan una configuración de seguridad para habilitar el cifrado en tránsito.

Versión 5.7.0

Las siguientes notas de la versión incluyen información sobre la versión 5.7.0 de Amazon EMR. Los cambios son respecto a la versión 5.6.0 de Amazon EMR.

Fecha de lanzamiento: 13 de julio de 2017

Actualizaciones

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

Nuevas características

- Se ha añadido la posibilidad de especificar una AMI de Amazon Linux personalizada al crear un clúster. Para más información, consulte Uso de una AMI personalizada.

Cambios, mejoras y problemas resueltos

- HBase
  - Se agregó la capacidad de configurar clústeres de HBase lectura y réplica. Consulte Uso de un clúster de réplicas de lectura.
  - Varias correcciones de errores y mejoras
- Presto: se ha agregado la capacidad de configurar `node.properties`.
- YARN: se ha agregado la capacidad de configurar `container-log4j.properties`
- Sqoop: se ha adaptado SQOOP-2880, que introduce un argumento que le permite definir el directorio temporal Sqoop.

Versión 5.6.0

Las siguientes notas de la versión incluyen información sobre la versión 5.6.0 de Amazon EMR. Los cambios son respecto a la versión 5.5.0 de Amazon EMR.

Fecha de lanzamiento: 5 de junio de 2017

Actualizaciones

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Esta es la primera versión de Mahout que admite Spark 2.x en la versión 5.0 y posteriores de Amazon EMR.
- Spark 2.1.1

Cambios, mejoras y problemas resueltos

- Presto

- Se agregó la capacidad de habilitar la comunicación SSL/TLS segura entre los nodos de Presto al habilitar el cifrado en tránsito mediante una configuración de seguridad. Para más información, consulte Cifrado de datos en tránsito.
- Se ha adaptado Presto 7661, que añade la opción `VERBOSE` a la instrucción `EXPLAIN ANALYZE` para notificar estadísticas de bajo nivel más detalladas sobre un plan de consulta.

Versión 5.5.3

Las siguientes notas de la versión incluyen información sobre la versión 5.5.3 de Amazon EMR. Los cambios son respecto a la versión 5.5.2.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- Esta versión aborda una posible vulnerabilidad de seguridad.

Versión 5.5.2

Las siguientes notas de la versión incluyen información sobre la versión 5.5.2 de Amazon EMR. Los cambios son respecto a la versión 5.5.1.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versión 5.5.1

Las siguientes notas de la versión incluyen información sobre la versión 5.5.1 de Amazon EMR. Los cambios son respecto a la versión 5.5.0 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versión 5.5.0

Las siguientes notas de la versión incluyen información sobre la versión 5.5.0 de Amazon EMR. Los cambios son respecto a la versión 5.4.0 de Amazon EMR.

Fecha de lanzamiento: 26 de abril de 2017

Actualizaciones

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1
- ZooKeeper 3.4.10

Cambios, mejoras y problemas resueltos

- Spark
  - Se ha corregido un parche de Spark (SPARK-20115) compatible con la versión DAGScheduler 2.1.0 de Spark, incluida en esta versión, que permitía volver a calcular todos los bloques de reproducción aleatoria perdidos cuando el servicio de reproducción aleatoria externo no estaba disponible.
- Flink
  - Flink ahora se compila con Scala 2.11. Si utiliza las bibliotecas y la API de Scala, le recomendamos que utilice Scala 2.11 en sus proyectos.
  - Se ha resuelto un problema donde los valores predeterminados de `HADOOP_CONF_DIR` y `YARN_CONF_DIR` no se definían correctamente, por tanto no funcionaba `start-scala-shell.sh`. También se ha añadido la capacidad de definir estos valores utilizando `env.hadoop.conf.dir` y `env.yarn.conf.dir` en `/etc/flink/conf/flink-conf.yaml` o la clasificación de configuración `flink-conf`.

  - Se ha introducido un nuevo comando específico de EMR, `flink-scala-shell` como encapsulador para `start-scala-shell.sh`. Recomendamos utilizar este comando en lugar de `start-scala-shell`. El nuevo comando simplifica la ejecución. Por ejemplo, `flink-scala-shell -n 2` inicia un shell Scala de Flink con un paralelismo de tareas de 2.
  - Se ha introducido un nuevo comando específico de EMR, `flink-yarn-session` como encapsulador para `yarn-session.sh`. Recomendamos utilizar este comando en lugar de `yarn-session`. El nuevo comando simplifica la ejecución. Por ejemplo, `flink-yarn-session -d -n 2` inicia una sesión de Flink de ejecución prolongada en un estado desasociado con dos administradores de tareas.
  - Se ha corregido (FLINK-6125) Commons httpclient is not shaded anymore in Flink 1.2.
- Presto
  - Se ha agregado compatibilidad para autenticación LDAP. El uso de LDAP con Presto en Amazon EMR requiere habilitar el acceso HTTPS para el coordinador de Presto (`http-server.https.enabled=true` en `config.properties`). Para conocer los detalles de configuración, consulte LDAP Authentication en la documentación de Presto.
  - Se agregó compatibilidad con `SHOW GRANTS`.
- AMI base de Linux de Amazon EMR
  - Las versiones de Amazon EMR ahora se basan en Amazon Linux 2017.03. Para más información, consulte las notas de la versión de Amazon Linux AMI 2017.03.
  - Se ha eliminado Python 2.6 de la imagen de Linux de base en Amazon EMR. Python 2.7 y 3.4 se instalan de forma predeterminada. Puede instalar Python 2.6 manualmente si es necesario.

Versión 5.4.0

Las siguientes notas de la versión incluyen información sobre la versión 5.4.0 de Amazon EMR. Los cambios son respecto a la versión 5.3.0 de Amazon EMR.

Fecha de lanzamiento: 8 de marzo de 2017

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a Flink 1.2.0
- Actualizado a Hbase 1.3.0
- Actualizado a Phoenix 4.9.0

> **Note**
>
> Si actualiza desde una versión anterior de Amazon EMR a una versión 5.4.0 o posterior de Amazon EMR y utiliza la indexación secundaria, actualice los índices locales como se describe en la documentación de Apache Phoenix. Amazon EMR elimina las configuraciones requeridas de la clasificación de `hbase-site`, pero los índices se tienen que volver a rellenar. Se admite la actualización de índices en línea y fuera de línea. Las actualizaciones en línea son la opción predeterminada, lo que significa que los índices se rellenan al inicializar desde los clientes de Phoenix versión 4.8.0 o superior. Para especificar actualizaciones fuera de línea, defina la configuración `phoenix.client.localIndexUpgrade` en false en la clasificación `phoenix-site` y, a continuación, SSH en el nodo principal para ejecutar `psql [zookeeper] -1`.

- Actualizado a Presto 0.166
- Actualizado a Zeppelin 0.7.0

Cambios y mejoras

Se han hechos los siguientes cambios en versiones de Amazon EMR para la etiqueta de versión emr-5.4.0:

- Se ha añadido compatibilidad para instancias r4. Consulta los tipos de EC2 instancias de Amazon.

Versión 5.3.1

Las siguientes notas de la versión incluyen información sobre la versión 5.3.1 de Amazon EMR. Los cambios son respecto a la versión 5.3.0 de Amazon EMR.

Fecha de lanzamiento: 7 de febrero de 2017

Cambios menores para agregar portabilidad con versiones anteriores de las revisiones de Zeppelin y actualizar la AMI predeterminada de Amazon EMR.

Versión 5.3.0

Las siguientes notas de la versión incluyen información sobre la versión 5.3.0 de Amazon EMR. Los cambios son respecto a la versión 5.2.1 de Amazon EMR.

Fecha de lanzamiento: 26 de enero de 2017

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a Hive 2.1.1
- Actualizado a Hue 3.11.0
- Actualizado a Spark 2.1.0
- Actualizado a Oozie 4.3.0
- Actualizado a Flink 1.1.4

Cambios y mejoras

Se han hechos los siguientes cambios en versiones de Amazon EMR para la etiqueta de versión emr-5.3.0:

- Se ha añadido un parche a Hue que le permite utilizar el ajuste `interpreters_shown_on_wheel` para configurar qué intérpretes se muestran primero en la rueda de selección de bloc de notas, independientemente de su orden en el archivo `hue.ini`.
- Se ha añadido la clasificación de configuración `hive-parquet-logging`, que se puede utilizar para configurar valores en el archivo `parquet-logging.properties` de Hive.

Versión 5.2.2

Las siguientes notas de la versión incluyen información sobre la versión 5.2.2 de Amazon EMR. Los cambios son respecto a la versión 5.2.1 de Amazon EMR.

Fecha de lanzamiento: 2 de mayo de 2017

Problemas conocidos resueltos de las versiones anteriores

- Se actualizó el SPARK-194459, que solucionó un problema por el que la lectura de una tabla ORC con columnas podía fallar. char/varchar

Versión 5.2.1

Las siguientes notas de la versión incluyen información sobre la versión 5.2.1 de Amazon EMR. Los cambios son respecto a la versión 5.2.0 de Amazon EMR.

Fecha de lanzamiento: 29 de diciembre de 2016

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Se ha actualizado Presto a la versión 0.157.1. Para más información, consulte las Presto Release Notes en la documentación de Presto.
- Se ha actualizado ZooKeeper a la versión 3.4.9. Para obtener más información, consulte las notas de la ZooKeeper versión en la documentación de Apache. ZooKeeper

Cambios y mejoras

Se han hechos los siguientes cambios en versiones de Amazon EMR para la etiqueta de versión emr-5.2.1:

- Se agregó compatibilidad con el tipo de instancia Amazon EC2 m4.16xlarge en la versión 4.8.3 y posteriores de Amazon EMR, excluidas las 5.0.0, 5.0.3 y 5.2.0.
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.
- La ubicación de las rutas de configuración de Flink y YARN ahora se define de forma predeterminada en `/etc/default/flink` y no es necesario definir las variables de entorno `FLINK_CONF_DIR` y `HADOOP_CONF_DIR` al ejecutar los scripts de controlador `flink` o `yarn-session.sh` para lanzar los trabajos de Flink.
- Se agregó soporte para FlinkKinesisConsumer la clase.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema en Hadoop que provocaba que el ReplicationMonitor hilo se quedara atascado durante mucho tiempo debido a una carrera entre la replicación y la eliminación del mismo archivo en un clúster grande.

- Se solucionó un problema por el que ControlledJob #toString fallaba con una excepción de puntero nulo (NPE) cuando el estado del trabajo no se actualizaba correctamente.

Versión 5.2.0

Las siguientes notas de la versión incluyen información sobre la versión 5.2.0 de Amazon EMR. Los cambios son respecto a la versión 5.1.0 de Amazon EMR.

Fecha de lanzamiento: 21 de noviembre de 2016

Cambios y mejoras

Los siguientes cambios y mejoras están disponibles en esta versión:

- Se agregó el modo de almacenamiento Amazon S3 para HBase.
- Le permite especificar una ubicación de Amazon S3 para el HBase directorio raíz. Para obtener más información, consulte HBaseAmazon S3.

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a Spark 2.0.2

Problemas conocidos resueltos de las versiones anteriores

- Corregido un error con /mnt limitado a 2 TB en tipos de instancias solo para EBS.
- Se ha corregido un error con registros instance-controller y logpusher que generan sus archivos .out correspondientes en lugar de los archivos .log configurados para log4j normales, que rotan cada hora. Los archivos .out no rotan, por lo que esto finalmente rellenaría la partición /emr. Este problema solo afecta a los tipos de instancia de máquina virtual de hardware (HVM).

Versión 5.1.0

Las siguientes notas de la versión incluyen información sobre la versión 5.1.0 de Amazon EMR. Los cambios son respecto a la versión 5.0.0 de Amazon EMR.

Fecha de lanzamiento: 3 de noviembre de 2016

Cambios y mejoras

Los siguientes cambios y mejoras están disponibles en esta versión:

- Se ha añadido soporte para Flink 1.1.3.
- Presto se ha añadido como una opción en la sección de bloc de notas de Hue.

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a la versión HBase 1.2.3
- Actualizado a Zeppelin 0.6.2

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema con las consultas de Tez en Amazon S3 con archivos de ORC que no tienen un rendimiento tan bueno como en las versiones de Amazon EMR anteriores a la 4.x.

Versión 5.0.3

Las siguientes notas de la versión incluyen información sobre la versión 5.0.3 de Amazon EMR. Los cambios son respecto a la versión 5.0.0 de Amazon EMR.

Fecha de lanzamiento: 24 de octubre de 2016

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a Hadoop 2.7.3
- Actualizado a Presto 0.152.3, que incluye soporte para la interfaz web de Presto. Puede acceder a la interfaz web de Presto en el coordinador de Presto utilizando el puerto 8889. Para más información acerca de la interfaz web de Presto, consulte Interfaz web en la documentación de Presto.

- Actualizado a Spark 2.0.1
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versión 5.0.0

Fecha de lanzamiento: 27 de julio de 2016

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a Hive 2.1
- Actualizado a Presto 0.150
- Actualizado a Spark 2.0
- Actualizado a Hue 3.10.0
- Actualizado a Pig 0.16.0
- Actualizado a Tez 0.8.4
- Actualizado a Zeppelin 0.6.1

Cambios y mejoras

Se han hechos los siguientes cambios en versiones de Amazon EMR para la etiqueta de versión emr-5.0.0 o superior:

- Amazon EMR es compatible con las últimas versiones de código abierto de Hive (versión 2.1) y Pig (versión 0.16.0). Si ha utilizado Hive o Pig en Amazon EMR en el pasado, esto podría afectar a algunos casos de uso. Para obtener más información, consulte Hive y Pig.
- El motor de ejecución predeterminado para Hive y Pig es ahora Tez. Para cambiarlo, debería editar los valores adecuados en las clasificaciones de configuración `hive-site` y `pig-properties`, respectivamente.
- Se ha añadido una característica de depuración de paso mejorada, que le permite ver la causa raíz de los errores de paso en caso de que el servicio pueda determinar la causa. Para más información, consulte Depuración escalonada mejorada en la Guía de administración de Amazon EMR.

- Las aplicaciones que anteriormente terminaban por "-Sandbox" ya no tienen ese sufijo. Esto podría interrumpir la automatización, por ejemplo, si utiliza scripts para lanzar clústeres con estas aplicaciones. En la siguiente tabla se muestran los nombres de aplicación la versión 4.7.2 de Amazon EMR en comparación con su versión 5.0.0.

  Cambios de nombre de aplicación

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-Sandbox	ZooKeeper

- Spark ya está compilado para Scala 2.11.
- Java 8 es ahora el JVM predeterminado. Todas las aplicaciones se ejecutan utilizando el tiempo de ejecución de Java 8. No hay cambios en el destino de código de bytes de ninguna aplicación. La mayoría de las aplicaciones siguen estableciendo como destino Java 7.
- Zeppelin ahora incluye características de autenticación. Para obtener más información, consulte Zeppelin.
- Se ha añadido soporte para configuraciones de seguridad, que le permiten crear y aplicar opciones de cifrado con más facilidad. Para más información, consulte Cifrado de datos.

Versión 4.9.5

Las siguientes notas de la versión incluyen información sobre la versión 4.9.5 de Amazon EMR. Los cambios son respecto a la versión 4.9.4.

Fecha de lanzamiento inicial: 29 de agosto de 2018

Cambios, mejoras y problemas resueltos

- HBase
  - Esta versión aborda una posible vulnerabilidad de seguridad.

Versión 4.9.4

Las siguientes notas de la versión incluyen información sobre la versión 4.9.4 de Amazon EMR. Los cambios son respecto a la versión 4.9.3.

Fecha de lanzamiento inicial: 29 de marzo de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar posibles vulnerabilidades.

Versión 4.9.3

Las siguientes notas de la versión incluyen información sobre la versión 4.9.3 de Amazon EMR. Los cambios son respecto a la versión 4.9.2 de Amazon EMR.

Fecha de lanzamiento inicial: 22 de enero de 2018

Cambios, mejoras y problemas resueltos

- Se ha actualizado el kernel de Amazon Linux de la AMI predeterminada de Amazon Linux para Amazon EMR a fin de abordar las vulnerabilidades asociadas a la ejecución especulativa (CVE-2017-5715, CVE-2017-5753 y CVE-2017-5754). Para obtener más información, consulte https://aws.amazon.com/security/security-bulletins/AWS-2018-013/.

Versión 4.9.2

Las siguientes notas de la versión incluyen información sobre la versión 4.9.2 de Amazon EMR. Los cambios son respecto a la versión 4.9.1 de Amazon EMR.

Fecha de lanzamiento: 13 de julio de 2017

Se han realizado pequeños cambios, correcciones de errores y mejoras en esta versión.

Versión 4.9.1

Las siguientes notas de la versión incluyen información sobre la versión 4.9.1 de Amazon EMR. Los cambios son respecto a la versión 4.8.4 de Amazon EMR.

Fecha de lanzamiento: 10 de abril de 2017

Problemas conocidos resueltos de las versiones anteriores

- Se han adaptado HIVE-9976 y HIVE-10106
- Se ha corregido un problema en YARN donde un gran número de nodos (superior a 2 000) y contenedores (superiores a 5 000) provocan un error de falta de memoria, por ejemplo: `"Exception in thread 'main' java.lang.OutOfMemoryError"`.

Cambios y mejoras

Se han hechos los siguientes cambios en versiones de Amazon EMR para la etiqueta de versión emr-4.9.1:

- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2017.03. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/.
- Se ha eliminado Python 2.6 de la imagen de Linux de base en Amazon EMR. Puede instalar Python 2.6 manualmente si es necesario.

Versión 4.8.4

Las siguientes notas de la versión incluyen información sobre la versión 4.8.4 de Amazon EMR. Los cambios son respecto a la versión 4.8.3 de Amazon EMR.

Fecha de lanzamiento: 7 de febrero de 2017

Se han realizado pequeños cambios, correcciones de errores y mejoras en esta versión.

Versión 4.8.3

Las siguientes notas de la versión incluyen información sobre la versión 4.8.3 de Amazon EMR. Los cambios son respecto a la versión 4.8.2 de Amazon EMR.

Fecha de lanzamiento: 29 de diciembre de 2016

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Se ha actualizado Presto a la versión 0.157.1. Para más información, consulte las Presto Release Notes en la documentación de Presto.

- Actualizado a Spark 1.6.3. Para más información, consulte las Spark Release Notes en la documentación de Apache Spark.
- Actualizado a 3.4.9. ZooKeeper Para obtener más información, consulte las notas ZooKeeper de la versión en la ZooKeeper documentación de Apache.

Cambios y mejoras

Se han hechos los siguientes cambios en versiones de Amazon EMR para la etiqueta de versión emr-4.8.3:

- Se agregó compatibilidad con el tipo de EC2 instancia Amazon m4.16xlarge en Amazon EMR versión 4.8.3 y posteriores, excepto 5.0.0, 5.0.3 y 5.2.0.
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema en Hadoop que provocaba que el ReplicationMonitor hilo se quedara atascado durante mucho tiempo debido a una carrera entre la replicación y la eliminación del mismo archivo en un clúster grande.
- Se solucionó un problema por el que ControlledJob #toString fallaba con una excepción de puntero nulo (NPE) cuando el estado del trabajo no se actualizaba correctamente.

Versión 4.8.2

Las siguientes notas de la versión incluyen información sobre la versión 4.8.2 de Amazon EMR. Los cambios son respecto a la versión 4.8.0 de Amazon EMR.

Fecha de lanzamiento: 24 de octubre de 2016

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a Hadoop 2.7.3
- Actualizado a Presto 0.152.3, que incluye soporte para la interfaz web de Presto. Puede acceder a la interfaz web de Presto en el coordinador de Presto utilizando el puerto 8889. Para más

información acerca de la interfaz web de Presto, consulte Interfaz web en la documentación de Presto.

- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.09. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/.

Versión 4.8.0

Fecha de lanzamiento: 7 de septiembre de 2016

Actualizaciones

Las siguientes actualizaciones están disponibles en esta versión:

- Actualizado a la versión HBase 1.2.2
- Actualizado a Presto-Sandbox 0.151
- Actualizado a Tez 0.8.4
- Actualizado a Zeppelin-Sandbox 0.6.1

Cambios y mejoras

Se han hechos los siguientes cambios en versiones de Amazon EMR para la etiqueta de versión emr-4.8.0:

- Se ha corregido un error en YARN que ApplicationMaster provocaba que intentaran limpiar contenedores que ya no existen porque sus instancias se habían cerrado.
- Se ha corregido la URL de hive-server2 para acciones de Hive2 en los ejemplos de Oozie.
- Se ha añadido soporte para catálogos de Presto adicionales.
- Se han adaptado los parches: HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- Se ha añadido soporte para configuraciones de seguridad, que le permiten crear y aplicar opciones de cifrado con más facilidad. Para más información, consulte Cifrado de datos.

Versión 4.7.2

Las siguientes notas de la versión incluyen información sobre la versión 4.7.2 de Amazon EMR.

Fecha de lanzamiento: 15 de julio de 2016

Características

Las siguientes características están disponibles en esta versión:

- Actualizado a Mahout 0.12.2
- Actualizado a Presto 0.148
- Actualizado a Spark 1.6.2
- Ahora puedes crear un AWSCredentials proveedor para usarlo con EMRFS utilizando un URI como parámetro. Para obtener más información, consulte Crear un AWSCredentials proveedor para EMRFS.
- EMRFS ahora permite a los usuarios configurar un punto de conexión de DynamoDB personalizado para sus metadatos de vista coherente utilizando la propiedad `fs.s3.consistent.dynamodb.endpoint` en `emrfs-site.xml`.
- Se ha añadido un script en `/usr/bin` denominado `spark-example`, que envuelve `/usr/lib/spark/spark/bin/run-example` para que pueda ejecutar los ejemplos directamente. Por ejemplo, para ejecutar el SparkPi ejemplo que viene con la distribución de Spark, puedes ejecutarlo `spark-example SparkPi 100` desde la línea de comandos o usarlo `command-runner.jar` como un paso de la API.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema donde Oozie no tenía `spark-assembly.jar` en la ubicación correcta cuando Spark también está instalado, lo que se traducía en un error al lanzar aplicaciones de Spark con Oozie.
- Se ha solucionado un problema con registros basado en Spark Log4j en contenedores de YARN.

Versión 4.7.1

Fecha de lanzamiento: 10 de junio de 2016

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un error que ampliaba el tiempo de startup de clústeres lanzadas en una VPC con subredes privadas. El error únicamente afectaba a clústeres lanzados con la versión 4.7.0 de Amazon EMR.

- Se ha corregido un error que gestionaba incorrectamente el listado de archivos en Amazon EMR para los clústeres iniciados con la versión 4.7.0 de Amazon EMR.

Versión 4.7.0

> **Important**
>
> Amazon EMR 4.7.0 está obsoleto. Utilice Amazon EMR 4.7.1 o una versión posterior en su lugar.

Fecha de lanzamiento: 2 de junio de 2016

Características

Las siguientes características están disponibles en esta versión:

- Se ha añadido Apache Phoenix 4.7.0
- Se ha añadido Apache Tez 0.8.3
- Actualizado a la versión HBase 1.2.1
- Actualizado a Mahout 0.12.0
- Actualizado a Presto 0.147
- Se actualizó AWS SDK para Java a la versión 1.10.75
- La marca final se ha eliminado de la propiedad `mapreduce.cluster.local.dir` en `mapred-site.xml` para permitir a los usuarios ejecutar Pig en modo local.

Los controladores JDBC de Amazon Redshift están disponibles en el clúster

Los controladores JDBC de Amazon Redshift ahora se incluyen en `/usr/share/aws/redshift/jdbc`. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar` es el controlador de Amazon Redshift compatible con JDBC 4.1 y `/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar` es el controlador de Amazon Redshift compatible con JDBC 4.0. Para más información, consulte Configurar una conexión JDBC en la Guía de administración de Amazon Redshift.

Java 8

Excepto para Presto, OpenJDK 1.7 es el JDK predeterminado para todas las aplicaciones. No obstante, tanto OpenJDK 1.7 como 1.8 están instalados. Para más información sobre cómo definir `JAVA_HOME` para las aplicaciones, consulte Configuración de aplicaciones para utilizar Java 8.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema de kernel que afectaba notablemente a los volúmenes de HDD de rendimiento optimizado (st1) de EBS para Amazon EMR en emr-4.6.0.
- Se ha corregido un problema donde un clúster fallaría si alguna de la zona de cifrado de HDFS se especifica sin elegir Hadoop como aplicación.
- Se ha cambiado la política de escritura de HDFS predeterminada `RoundRobin` a `AvailableSpaceVolumeChoosingPolicy`. Algunos volúmenes no se utilizaron correctamente con la RoundRobin configuración, lo que provocó fallos en los nodos principales y un HDFS poco fiable.
- Se ha corregido un problema con la CLI de EMRFS que provocaría una excepción al crear la tabla de metadatos de DynamoDB predeterminada para vistas coherentes.
- Se ha corregido un problema de bloqueo en EMRFS que ocurría potencialmente durante las operaciones de copia y cambio de nombre de multipartes.
- Se ha corregido un problema con EMRFS que provocaba que el CopyPart tamaño predeterminado fuera de 5 MB. El valor predeterminado se ha fijado correctamente ahora en 128 MB.
- Se ha corregido un problema con la configuración de upstart de Zeppelin que potencialmente impedía parar el servicio.
- Se ha corregido un problema con Spark y Zeppelin, que impedía que utilizara el esquema de URI de `s3a://` porque `/usr/lib/hadoop/hadoop-aws.jar` no se cargó correctamente en la classpath respectiva.
- Se ha adaptado HUE-2484.
- Se respaldó una confirmación de Hue 3.9.0 (no existe JIRA) para solucionar un problema con la muestra del navegador. HBase
- Se ha adaptado HIVE-9073.

Versión 4.6.0

Fecha de lanzamiento: 21 de abril de 2016

Características

Las siguientes características están disponibles en esta versión:

- HBase Se agregó la versión 1.2.0
- Se ha añadido Zookeeper-Sandbox 3.4.8
- Actualizado a Presto-Sandbox 0.143
- Las versiones de Amazon EMR ahora se basan en Amazon Linux 2016.03.0. Para obtener más información, consulte https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/.

Problema que afecta a los tipos de volúmenes de HDD de rendimiento optimizado (st1) de EBS

Un problema en el kernel de Linux versiones 4.2 y superiores afecta notablemente al rendimiento en volúmenes de EBS HDD con velocidad optimizada (st1) para EMR. Esta versión (emr-4.6.0) utiliza el kernel versión 4.4.5 y, por tanto, se ve afectada. Por lo tanto, le recomendamos no utilizar emr-4.6.0 si desea utilizar los volúmenes de EBS st1. Puede utilizar emr-4.5.0 o versiones anteriores de Amazon EMR sin que afecten a st1. Además, proporcionamos la corrección en futuras versiones.

Valores predeterminados de Python

Python 3.4 se instala ahora de forma predeterminada, pero Python 2.7 sigue siendo el valor predeterminado del sistema. Puede configurar Python 3.4 como predeterminado del sistema mediante una acción de arranque; puede usar la API de configuración para establecer la exportación de PYSPARK_PYTHON `/usr/bin/python3.4` en la `spark-env` clasificación para que afecte a la versión de Python utilizada por. PySpark

Java 8

Excepto para Presto, OpenJDK 1.7 es el JDK predeterminado para todas las aplicaciones. No obstante, tanto OpenJDK 1.7 como 1.8 están instalados. Para más información sobre cómo definir `JAVA_HOME` para las aplicaciones, consulte Configuración de aplicaciones para utilizar Java 8.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema donde el aprovisionamiento de aplicaciones en ocasiones fallaría aleatoriamente debido a una contraseña generada.

- Anteriormente, `mysqld` estaba instalado en todos los nodos. Ahora, solo está instalado en la instancia principal y únicamente si la aplicación elegida incluye `mysql-server` como componente. Actualmente, las siguientes aplicaciones incluyen el `mysql-server` componente: Hive, Hue HCatalog, Presto-Sandbox y Sqoop-Sandbox.
- Se ha cambiado `yarn.scheduler.maximum-allocation-vcores` a 80 desde el valor predeterminado de 32, que corrige un problema introducido en emr-4.4.0 que se produce sobre todo con Spark al utilizar la opción `maximizeResourceAllocation` en un clúster cuyo tipo de instancia secundaria es uno de los pocos tipos de instancia grandes que tienen los vcores de YARN definidos en un valor superior a 32; es decir, c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge o m4.10xlarge se han visto afectados por este problema.
- s3-dist-cp ahora utiliza EMRFS para todas las nominaciones de Amazon S3 y ya no hay fases en un directorio temporal de HDFS.
- Se ha corregido un problema con la gestión de excepciones para las cargas multiparte de cifrado del lado del cliente.
- Se ha agregado una opción para permitir a los usuarios cambiar la clase de almacenamiento de Amazon S3. De forma predeterminada, este valor es `STANDARD`. El valor de clasificación de configuración `emrfs-site` es `fs.s3.storageClass` y los valores posibles son `STANDARD`, `STANDARD_IA` y `REDUCED_REDUNDANCY`. Para más información sobre las clases de almacenamiento, consulte Clases de almacenamiento en la Guía del usuario de Amazon Simple Storage Service.

Versión 4.5.0

Fecha de lanzamiento: 4 de abril de 2016

Características

Las siguientes características están disponibles en esta versión:

- Actualizado a Spark 1.6.1
- Actualizado a Hadoop 2.7.2
- Actualizado a Presto 0.140
- Se agregó AWS KMS compatibilidad con el cifrado del lado del servidor de Amazon S3.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema donde los servidores de MySQL y Apache no se iniciaban después de reiniciar un nodo.
- Se ha corregido un problema donde IMPORT no funcionaba correctamente con tablas no particionadas almacenadas en Amazon S3
- Se ha corregido un problema en Presto donde requiere que el directorio intermedio sea `/mnt/tmp` en lugar de `/tmp` al escribir en tablas de Hive.

Versión 4.4.0

Fecha de lanzamiento: 14 de marzo de 2016

Características

Las siguientes características están disponibles en esta versión:

- Se agregó la versión 1.0.0 HCatalog
- Se ha añadido Sqoop-Sandbox 1.4.6
- Actualizado a Presto 0.136
- Actualizado a Zeppelin 0.5.6
- Actualizado a Mahout 0.11.1
- `dynamicResourceAllocation` habilitado de forma predeterminada.
- Se ha añadido una tabla de todas las clasificaciones de configuración para esta versión. Para más información, consulte la tabla de clasificaciones de configuración de Configuración de aplicaciones.

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un error que provocaba que la `maximizeResourceAllocation` configuración no reservara suficiente memoria para los demonios de YARN ApplicationMaster.
- Se ha corregido un problema encontrado con una DNS personalizada. Si alguna entrada en `resolve.conf` precede a las entradas personalizadas proporcionada, entonces las entradas personalizadas no se pueden resolver. Este comportamiento se veía afectado por clústeres en una VPC donde el servidor de nombres de VPC predeterminado se inserta como entrada superior en `resolve.conf`.

- Se ha corregido un problema donde la versión predeterminada de Python se traslada a la versión 2.7 y boto no se ha instalado para dicha versión.
- Se ha corregido un error donde los contenedores de YARN y las aplicaciones de Spark generarían un archivo único de base de datos round robin (rrd) de Ganglia, que daría lugar a que se llenara el primer disco asociado a la instancia. Debido a esta solución, las métricas en el nivel de contenedor de YARN se han deshabilitado y las métricas en el nivel de aplicación de Spark se han deshabilitado.
- Se ha corregido un problema en el insertador de registros que eliminaría todas las carpetas con registro vacío. El resultado era que la CLI de Hive no podía realizar el registro porque el insertador de registros eliminaba la carpeta `user` vacía en `/var/log/hive`.
- Se ha corregido un problema que afectaba a las importaciones de Hive, que afectaba a las particiones y daba lugar a un error durante la importación.
- Se ha corregido un problema donde EMRFS y s3-dist-cp no gestionaban de forma adecuada los nombres de bucket que contenían puntos.
- Se ha cambiado un comportamiento en EMRFS, de forma que los buckets con control de versiones habilitado en el archivo de marcador `_$folder$` no se creaba de forma continua, lo que podría contribuir a mejorar el rendimiento de los buckets con control de versiones habilitado.
- Se ha cambiado el comportamiento en EMRFS de forma que no utiliza archivos de instrucciones, excepto en aquellos casos en los que el cifrado del lado del cliente está habilitado. Si desea eliminar archivos de instrucciones al utilizar el cifrado del lado cliente, puede establecer la propiedad de emrfs-site.xml, `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled`, en true.
- Se ha cambiado la agregación de registros de YARN para conservar los registros en el destino agregación durante dos días. El destino predeterminado es el almacenamiento de HDFS de su clúster. Si desea cambiar esta duración, cambie el valor de `yarn.log-aggregation.retain-seconds` utilizando la clasificación de configuración `yarn-site` al crear el clúster. Como siempre, puede guardar los registros de las aplicaciones en Amazon S3 utilizando el parámetro `log-uri` al crear el clúster.

Parches aplicados

En esta versión se han incluido los siguientes parches de proyectos de código abierto:

- HIVE-9655
- HIVE-9183

- HADOOP-12810

Versión 4.3.0

Fecha de lanzamiento: 19 de enero de 2016

Características

Las siguientes características están disponibles en esta versión:

- Actualizado a Hadoop 2.7.1
- Actualizado a Spark 1.6.0
- Actualizado Ganglia a 3.7.2
- Actualizado Presto a 0.130

Amazon EMR ha hecho algunos cambios en `spark.dynamicAllocation.enabled` cuando se define en true; su valor predeterminado es false. Cuando se define en true, esto afecta a los valores predeterminados del ajuste `maximizeResourceAllocation`:

- Si `spark.dynamicAllocation.enabled` se ha definido en true, `spark.executor.instances` no ha sido definido por `maximizeResourceAllocation`.
- El ajuste `spark.driver.memory` se configura ahora en función de los tipos de instancia en el clúster de forma similar a cómo se define `spark.executors.memory`. Sin embargo, dado que la aplicación del controlador Spark puede ejecutarse en la instancia principal o en una de las instancias secundarias (por ejemplo, en un cliente de YARN y modos de clústeres, respectivamente), el ajuste `spark.driver.memory` se basa en el tipo de instancia del tipo de instancia más pequeño de estos dos grupos de instancias.
- El ajuste `spark.default.parallelism` se define ahora como el doble del número de núcleos de CPU disponibles para contenedores de YARN. En las versiones anteriores, era la mitad de dicho valor.
- Los cálculos de la sobrecarga de memoria reservada para procesos de Spark YARN se ha ajustado para ser más precisa, lo que se traduce en un pequeño aumento en la cantidad total de memoria disponible para Spark (es decir, `spark.executor.memory`).

Problemas conocidos resueltos de las versiones anteriores

- La agregación de registros de YARN ahora está habilitada de forma predeterminada.
- Se ha corregido un problema por el que los registros no se enviaban a un bucket de registros de Amazon S3 del clúster cuando la agregación de registros de YARN estaba habilitada.
- Los tamaños de contenedor de YARN ahora tienen un nuevo mínimo de 32 en todos los tipos de nodos.
- Se ha corregido un problema con Ganglia que provocaba un exceso de disco I/O en el nodo principal en clústeres grandes.
- Se ha corregido un problema que impedía que los registros de las aplicaciones se enviaran a Amazon S3 cuando se cierra un clúster.
- Se ha corregido un problema en la CLI de EMRFS que provocaba que determinados comandos devolvieran error.
- Se ha corregido un problema con Zeppelin que impedía que las dependencias se cargaran en la base. SparkContext
- Se ha corregido un problema que se producía al intentar un ajuste de tamaño para añadir instancias.
- Se ha corregido un problema en Hive donde CREATE TABLE AS SELECT realiza llamadas de lista excesivas a Amazon S3.
- Se ha corregido un problema donde clústeres grandes no aprovisionarían correctamente cuando estaban instalados Hue, Oozie y Ganglia.
- Se ha corregido un problema en s3-dist-cp que devolvería un código de salida cero incluso si ha devuelto un error.

Parches aplicados

En esta versión se han incluido los siguientes parches de proyectos de código abierto:

- OOZIE-2402
- HIVE-12502
- HIVE-10631
- HIVE-12213
- HIVE-10559
- HIVE-12715

- HIVE-10685

Versión 4.2.0

Fecha de lanzamiento: 18 de noviembre de 2015

Características

Las siguientes características están disponibles en esta versión:

- Se ha añadido compatibilidad con Ganglia
- Actualizado a Spark 1.5.2
- Actualizado a Presto 0.125
- Actualizado Oozie a 4.2.0
- Actualizado Zeppelin a 0.5.5
- Se actualizó a la versión 1.10.27 AWS SDK para Java

Problemas conocidos resueltos de las versiones anteriores

- Se ha corregido un problema con la CLI de EMRFS, donde no se utilizaba el nombre de la tabla de metadatos predeterminado.
- Se ha corregido un problema que se presentaba al utilizar tablas respaldadas por ORC en Amazon S3.
- Se ha corregido un problema encontrado con un error de coincidencia de versión de Python en la configuración de Spark.
- Se ha corregido un error cuando falla la notificación del estado de un nodo de YARN debido a problemas de DNS para clústeres en una VPC.
- Se ha corregido un error que se encuentra cuando YARN retira nodos, lo que da lugar a aplicaciones que no responden o a la incapacidad de programar nuevas aplicaciones.
- Se ha corregido un error que se encuentra cuando los clústeres terminan con el estado TIMED_OUT_STARTING.
- Se ha corregido un error que se encuentra al incluir la dependencia de EMRFS Scala en otras versiones. Se ha eliminado la dependencia de Scala.

Configuración de aplicaciones

Para anular las configuraciones predeterminadas de una aplicación, puede subministrar un objeto de configuración. Puede utilizar tanto una sintaxis abreviada para proporcionar la configuración como hacer referencia al objeto de configuración en un archivo JSON. Los objetos de configuración se componen de una clasificación, propiedades y configuraciones anidadas opcionales. Las propiedades corresponden a la configuración de la aplicación que desee cambiar. Es posible especificar varias clasificaciones para varias aplicaciones en un solo objeto JSON.

> **Warning**
>
> Las operaciones de la API de descripción y lista de Amazon EMR emiten configuraciones personalizadas, que se pueden establecer y que se utilizan como parte de los flujos de trabajo de Amazon EMR, en texto sin formato. Para incluir información confidencial, como contraseñas, en estas configuraciones, consulte Almacenamiento de datos de configuración confidenciales en AWS Secrets Manager.

Las clasificaciones de configuración disponibles varían según la versión de Amazon EMR. Para obtener una lista de las clasificaciones de configuración compatibles con una versión de lanzamiento concreta, consulte la página correspondiente a esa versión en la sección Acerca de las versiones de Amazon EMR.

A continuación se muestra un ejemplo de archivo de JSON para obtener una lista de configuraciones.

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.security.groups.cache.secs": "250"
    }
  },
  {
    "Classification": "mapred-site",
    "Properties": {
      "mapred.tasktracker.map.tasks.maximum": "2",
      "mapreduce.map.sort.spill.percent": "0.90",
```

```
      "mapreduce.tasktracker.reduce.tasks.maximum": "5"
    }
  }
]
```

Una clasificación de configuración suele corresponderse con un archivo de configuración específico de la aplicación. Por ejemplo, la clasificación `hive-site` se corresponde con los valores del archivo de configuración `hive-site.xml` de Hive. Una excepción es la acción de arranque `configure-daemons` ya no admitida, que se utiliza para establecer parámetros de entorno como `--namenode-heap-size`. Las opciones de este tipo se engloban en las clasificaciones `hadoop-env` y `yarn-env` con sus propias clasificaciones de exportación anidadas. Si alguna clasificación termina por `env`, debe utilizar la subclasificación de exportación.

Otra excepción es `s3get`, que se utiliza para colocar un objeto `EncryptionMaterialsProvider` del cliente en cada nodo de un clúster para su uso en cifrado del cliente. Se ha añadido una opción a la clasificación `emrfs-site` para este fin.

A continuación se muestra un ejemplo de la clasificación `hadoop-env`.

```
[
  {
    "Classification": "hadoop-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "2048",
          "HADOOP_NAMENODE_OPTS": "-XX:GCTimeRatio=19"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

A continuación se muestra un ejemplo de la clasificación yarn-env.

```
[
  {
    "Classification": "yarn-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "YARN_RESOURCEMANAGER_OPTS": "-Xdebug -Xrunjdwp:transport=dt_socket"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

Los siguientes ajustes no pertenecen a un archivo de configuración, pero los utiliza Amazon EMR para configurar potencialmente diversos ajustes en su nombre.

Ajustes mantenidos por Amazon EMR

Aplicación	Clasificación de etiqueta de la versión	Propiedades válidas	Cuándo se debe usar
Spark	`spark`	`maximizeResourceAllocation`	Configure ejecutores para utilizar los recursos máximos de cada nodo.

Temas

- Configuración de las aplicaciones al crear un clúster
- Reconfiguración de un grupo de instancias en un clúster en ejecución
- Almacenamiento de datos de configuración confidenciales en AWS Secrets Manager
- Configuración de aplicaciones para utilizar una máquina virtual de Java específica

Configuración de las aplicaciones al crear un clúster

Al crear un clúster, puede anular las configuraciones predeterminadas para las aplicaciones mediante la consola de Amazon EMR,AWS CLI() o AWS Command Line Interface AWS el SDK de.

Para anular la configuración predeterminada de una aplicación, debe especificar valores personalizados en una clasificación de configuración. Una clasificación de configuración corresponde a un archivo XML de configuración para una aplicación, como `hive-site.xml`.

Las clasificaciones de configuración varían en función de la versión de Amazon EMR. Para ver una lista de las clasificaciones de configuración que están disponibles para cada versión de lanzamiento, consulte la página de detalles de lanzamiento. Por ejemplo, Versión 6.4.0. de Amazon EMR.

Proporcionar una configuración en la consola al crear un clúster

Para proporcionar una configuración, vaya a la página Crear clúster y amplíe la Configuración de software. A continuación, puede introducir la configuración directamente a través de JSON o una sintaxis abreviada mostrada en el texto en sombra en la consola. De lo contrario, puede proporcionar una URI de Amazon S3 para un archivo con un objeto `Configurations` de JSON.

Para proporcionar una configuración para un grupo de instancias, elija un clúster de la lista de clústeres y, a continuación, elija la pestaña Configuraciones. En la tabla de Configuraciones de grupos de instancias, seleccione el grupo de instancias que desea editar y, a continuación, seleccione Reconfigurar.

Proporcione una configuración mediante AWS CLI al crear un clúster

Puede proporcionar una configuración para create-cluster suministrando una ruta a un archivo JSON almacenado localmente o en Amazon S3. En el siguiente ejemplo se supone que está utilizando los roles predeterminados de Amazon EMR y que roles ya se han creado. Si necesita crear los roles, ejecute `aws emr create-default-roles` en primer lugar.

Si la configuración está en su directorio local, puede utilizar el siguiente comando de ejemplo.

```
aws emr create-cluster --use-default-roles --release-label emr-7.9.0 --applications
 Name=Hive \
--instance-type m5.xlarge --instance-count 3 --configurations file://./
configurations.json
```

Si su configuración está en una ruta de Amazon S3, tendrá que configurar la siguiente solución alternativa antes de pasar la ruta de Amazon S3 al comando `create-cluster`.

```
#!/bin/sh
# Assume the ConfigurationS3Path is not public, and its present in the same AWS account
 as the EMR cluster
ConfigurationS3Path="s3://amzn-s3-demo-bucket/config.json"
# Get a presigned HTTP URL for the s3Path
ConfigurationURL=`aws s3 presign $ConfigurationS3Path --expires-in 300`
# Fetch the presigned URL, and minify the JSON so that it spans only a single line
Configurations=`curl $ConfigurationURL | jq -c .`
aws emr create-cluster --use-default-roles --release-label emr-5.34.0 --instance-type
 m5.xlarge --instance-count 2 --applications Name=Hadoop Name=Spark --configurations
 $Configurations
```

Proporcione una configuración mediante el SDK de Java al crear un clúster

El siguiente extracto de programa muestra cómo proporcionar una configuración utilizando AWS SDK para Java.

```
Application hive = new Application().withName("Hive");

Map<String,String> hiveProperties = new HashMap<String,String>();
 hiveProperties.put("hive.join.emit.interval","1000");
 hiveProperties.put("hive.merge.mapfiles","true");

Configuration myHiveConfig = new Configuration()
 .withClassification("hive-site")
 .withProperties(hiveProperties);

RunJobFlowRequest request = new RunJobFlowRequest()
 .withName("Create cluster with ReleaseLabel")
 .withReleaseLabel("emr-5.20.0")
 .withApplications(hive)
 .withConfigurations(myHiveConfig)
 .withServiceRole("EMR_DefaultRole")
 .withJobFlowRole("EMR_EC2_DefaultRole")
 .withInstances(new JobFlowInstancesConfig()
  .withEc2KeyName("myEc2Key")
  .withInstanceCount(3)
  .withKeepJobFlowAliveWhenNoSteps(true)
  .withMasterInstanceType("m4.large")
```

```
  .withSlaveInstanceType("m4.large")
);
```

Reconfiguración de un grupo de instancias en un clúster en ejecución

Con la versión 5.21.0 y posteriores de Amazon EMR, puede volver a configurar las aplicaciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, puede utilizar la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK.

Cuando actualiza la configuración de una aplicación para un grupo de instancias en la nueva consola de Amazon EMR, la consola intenta combinar la nueva configuración con la configuración existente para crear una nueva configuración activa. En el caso excepcional de que Amazon EMR no pueda combinar la configuración, la consola le avisará.

Después de enviar una solicitud de reconfiguración para un grupo de instancias, Amazon EMR asigna un número de versión para la nueva especificación de configuración. Puede realizar un seguimiento del número de versión de una configuración o el estado de un grupo de instancias visualizando los CloudWatch eventos. Para obtener más información, consulta Supervisar CloudWatch eventos.

> **Note**
>
> Solo puede anular, pero no eliminar, las configuraciones de clúster que se hayan especificado durante la su creación. Si hay diferencias entre la configuración existente y el archivo que suministra, Amazon EMR restablece las configuraciones modificadas manualmente, como las que ha modificado mientras estaba conectado al clúster mediante SSH, a los valores predeterminados del clúster para el grupo de instancias especificado.

Consideraciones a la hora de volver a configurar un grupo de instancias

Acciones de reconfiguración

Cuando envía una solicitud de reconfiguración mediante la consola de Amazon EMR, AWS Command Line Interface AWS CLI() o SDK, Amazon EMR comprueba AWS el archivo de

configuración existente en el clúster. Si hay diferencias entre la configuración existente y el archivo que suministra, Amazon EMR inicia las acciones de reconfiguración, reinicia algunas aplicaciones y restablece cualquiera de las configuraciones modificadas manualmente, como las que ha modificado mientras estaba conectado al clúster mediante SSH, a los valores predeterminados del clúster para el grupo de instancias especificado.

> **Note**
>
> Amazon EMR realiza algunas acciones predeterminadas durante la reconfiguración de cada grupo de instancias. Estas acciones predeterminadas pueden entrar en conflicto con las personalizaciones del clúster que haya realizado y provocar errores de reconfiguración. Para obtener información sobre cómo solucionar errores de reconfiguración, consulte Solución de problemas de reconfiguración de grupos de instancias.

Amazon EMR también inicia acciones de reconfiguración para las clasificaciones de configuración que especifique en su solicitud. Para obtener una lista completa de estas acciones, consulte la sección Clasificaciones de configuración de la versión de Amazon EMR que utilice. Por ejemplo, Clasificaciones de configuración 6.2.0.

> **Note**
>
> En la guía de versiones de Amazon EMR solo se enumeran las acciones de reconfiguración a partir de las versiones 5.32.0 y 6.2.0 de Amazon EMR.

Interrupción del servicio

Amazon EMR sigue un proceso continuo para volver a configurar las instancias de los grupos de instancias Task y Core. Solo el 10 por ciento de las instancias de un grupo de instancias se modifican y se reinicia a la vez. Este proceso tarda más en terminar pero reduce la posibilidad de error de la aplicación en un clúster en ejecución.

Para ejecutar trabajos de YARN durante un reinicio de YARN, puede crear un clúster de Amazon EMR con varios nodos maestros o establecer `yarn.resourcemanager.recovery.enabled` en `true` en su clasificación de configuración `yarn-site`. Para obtener más información sobre el uso de varios nodos maestros, consulte YARN de alta disponibilidad ResourceManager.

Validación de aplicaciones

Amazon EMR comprueba que todas las aplicaciones del clúster se estén ejecutando tras el proceso de reinicio de la reconfiguración. Si alguna aplicación no está disponible, se produce un error en la operación de reconfiguración general. Si una reconfiguración da error, Amazon EMR invierte los parámetros de configuración a la versión anterior que funcionaba.

> **Note**
>
> Para evitar un error de reconfiguración, le recomendamos que solo instale en el clúster las aplicaciones que tenga pensado utilizar. También le recomendamos que se asegure de que todas las aplicaciones del clúster estén en buen estado y en funcionamiento antes de enviar una solicitud de reconfiguración.

Tipos de reconfiguración

Puede volver a configurar un grupo de instancias de una de estas dos maneras:

- Overwrite. Método de reconfiguración predeterminado y el único disponible en las versiones de Amazon EMR anteriores a las 5.35.0 y 6.6.0. Este método de reconfiguración sobrescribe indiscriminadamente todos los archivos del clúster con el conjunto de configuraciones recién enviado. El método borra cualquier cambio en los archivos de configuración realizado fuera de la API de reconfiguración.
- Merge. Método de reconfiguración compatible con las versiones 5.35.0 y 6.6.0 y posteriores de Amazon EMR, excepto desde la consola de Amazon EMR, que no admite ninguna versión. Este método de reconfiguración combina las configuraciones recién enviadas con las configuraciones que ya existen en el clúster. Esta opción solo agrega o modifica las nuevas configuraciones que envíe. Conserva las configuraciones existentes.

> **Note**
>
> Amazon EMR sigue sobrescribiendo algunas configuraciones esenciales de Hadoop que necesita para garantizar que el servicio se ejecute correctamente.

Limitaciones

Cuando vuelva a configurar un grupo de instancias en un clúster en ejecución, tenga en cuenta las siguientes limitaciones:

- Las aplicaciones que no son de YARN pueden fallar durante el reinicio o provocar problemas con el clúster, especialmente si no están configuradas correctamente. Los clústeres que se acerquen al uso máximo de memoria y CPU pueden tener problemas tras el proceso de reinicio. Esto es especialmente cierto en el caso del grupo de instancias maestro.
- No puede enviar una solicitud de reconfiguración cuando se está cambiando el tamaño de un grupo de instancias. Si una reconfiguración mientras se inicia un grupo de instancias se está cambiando de reconfiguración, no puede comenzar hasta que se haya completado el grupo de instancias de tamaño y viceversa.
- Después de configurar un grupo de instancias, Amazon EMR reinicia las aplicaciones para permitir que las nuevas configuraciones surtan efecto. Puede darse un error de trabajo u otro comportamiento de la aplicación inesperado si las aplicaciones están en uso durante la reconfiguración.
- Si una reconfiguración para un grupo de instancias da error, Amazon EMR invierte los parámetros de configuración a la versión anterior que funcionaba. Si el proceso de reversión también da error, debe enviar una nueva solicitud `ModifyInstanceGroup` para recuperar el grupo de instancias desde el estado `SUSPENDED`.
- Las solicitudes de reconfiguración para cualquier Phoenix clasificaciones de configuración solo se admiten en la versión 5.23.0 y versiones posteriores de Amazon EMR, y no son compatibles con la versión 5.21.0 o 5.22.0 de Amazon EMR.
- Las solicitudes de reconfiguración para cualquier HBase configuración clasificaciones de configuración solo se admiten en la versión 5.30.0 y versiones posteriores de Amazon EMR, y no son compatibles desde la versión 5.23.0 hasta la 5.29.0 de Amazon EMR.
- Amazon EMR admite solicitudes de reconfiguración de aplicaciones en un clúster de Amazon EMR con varios nodos principales solo en las versiones 5.27.0 y posteriores de Amazon EMR.
- La reconfiguración de la clasificación `hdfs-encryption-zones` o cualquiera de las clasificaciones de configuración de Hadoop KMS no se admite en clúster de Amazon EMR con varios nodos principales.
- Amazon EMR no admite actualmente determinadas solicitudes de reconfiguración para el programador de capacidad que requiere reiniciar el YARN. ResourceManager Por ejemplo, no puede eliminar por completo una cola.

Reconfiguración de un grupo de instancias en la consola

> **Note**
>
> La consola de Amazon EMR no admite reconfiguraciones de tipo Merge.

1. Abra la consola de Amazon EMR en /emr https://console.aws.amazon.com
2. En la lista de clústeres llamada Nombre, seleccione el clúster activo que desea volver a configurar.
3. Abra la página de detalles del clúster y vaya a la pestaña Configuraciones.
4. En la lista desplegable Filter, seleccione el grupo de instancias que desea volver a configurar.
5. En el menú desplegable Volver a configurar, elija Editar en la tabla o Editar en el archivo JSON.

   - Editar en la tabla: en la tabla de clasificación de configuración, edite la propiedad y el valor de las configuraciones existentes o elija Agregar configuración para suministrar clasificaciones de configuración adicionales.
   - Editar en el archivo JSON: ingrese la configuración directamente en el JSON o use la sintaxis abreviada (mostrada en el texto en sombra). De lo contrario, proporcione una URI de Amazon S3 para un archivo con un objeto `Configurations` de JSON.

   > **Note**
   >
   > La columna Origen de la tabla de clasificación de configuración indica si la configuración se proporciona al crear un clúster o cuando se especifican configuraciones adicionales para este grupo de instancias. Puede editar las configuraciones para un grupo de instancias de ambas fuentes. No puede eliminar las configuraciones de clúster iniciales, pero puede anularlas para un grupo de instancias.
   > También puede añadir o editar clasificaciones de configuración anidadas directamente a la tabla. Por ejemplo, para indicar una clasificación de subredes `export` adicionales `hadoop-env`, añada una clasificación de configuración `hadoop.export` a la tabla. A continuación, proporcione una propiedad específica y el valor de esta clasificación.
6. (Opcional) Seleccione Apply this configuration to all active instance groups (Aplicar esta configuración a todos los grupos de instancias activas).
7. Guarde los cambios.

Reconfiguración de un grupo de instancias con la CLI

Use el comando modify-instance-groups para especificar una nueva configuración para cada grupo de instancias en un clúster en ejecución.

> **Note**
>
> En los siguientes ejemplos, *<j-2AL4XXXXXX5T9>* sustituya por su ID de clúster y *<ig-1xxxxxxx9>* sustituya por su ID del grupo de instancias.

Example – Sustitución de una configuración para un grupo de instancias

En el siguiente ejemplo se hace referencia a un archivo de configuración JSON llamado `instanceGroups.json` para editar la propiedad del NodeManager disco comprobador de estado YARN para un grupo de instancias.

1. Prepare la clasificación de la configuración y guárdela como `instanceGroups.json` en el mismo directorio en el que ejecutará el comando.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Ejecute el siguiente comando.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example – Agregado de una configuración a un grupo de instancias

Si desea agregar una configuración a un grupo de instancias, debe incluir todas las configuraciones antes especificadas para ese grupo de instancias en su nueva solicitud `ModifyInstanceGroup`. De lo contrario, las configuraciones antes especificadas se eliminan.

En el siguiente ejemplo se agrega una propiedad para el comprobador de memoria NodeManager virtual de YARN. La configuración también incluye valores especificados previamente para el comprobador de estado del NodeManager disco de YARN para que los valores no se sobrescriban.

1. Prepare los contenidos siguientes en `instanceGroups.json` y guárdelos en el mismo directorio en el que va a ejecutará el comando.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0",
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Ejecute el siguiente comando.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
```

```
--instance-groups file://instanceGroups.json
```

Example – Agregar una configuración a un grupo de instancias con la reconfiguración de tipo Merge

Si desea usar el método de reconfiguración Overwrite predeterminado para agregar una configuración, debe incluir todas las configuraciones especificadas previamente para ese grupo de instancias en su nueva solicitud `ModifyInstanceGroup`. De lo contrario, el método Overwrite elimina las configuraciones que haya especificado anteriormente. No necesita hacer esto con la reconfiguración Merge. En su lugar, debe asegurarse de que su solicitud solo incluya las nuevas configuraciones.

En el siguiente ejemplo se agrega una propiedad para el comprobador de memoria NodeManager virtual de YARN. Como se trata de una reconfiguración de tipo Merge, no sobrescribe los valores especificados previamente para el comprobador de estado del NodeManager disco de YARN.

1. Prepare los contenidos siguientes en `instanceGroups.json` y guárdelos en el mismo directorio en el que va a ejecutará el comando.

```
[
   {"InstanceGroupId":"<ig-1xxxxxxx9>",
    "ReconfigurationType" :"MERGE",
      "Configurations":[
         {"Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Ejecute el siguiente comando.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example – Eliminación de una configuración para un grupo de instancias

Para eliminar una configuración para un grupo de instancias, envíe una nueva solicitud de reconfiguración que excluya la configuración anterior.

> **Note**
>
> Solo puede anular la configuración inicial del clúster. No puede eliminarla.

Por ejemplo, para eliminar la configuración del NodeManager disco comprobador de estado YARN que se especificó en el ejemplo anterior, puede enviar una nueva `instanceGroups.json` con el siguiente contenido.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

> **Note**
>
> Para eliminar todas las configuraciones de la última solicitud de reconfiguración, envíe una solicitud de reconfiguración con una matriz vacía de configuraciones. Por ejemplo:
>
> ```
> [
>    {
>       "InstanceGroupId":"<ig-1xxxxxxx9>",
>       "Configurations":[]
> ```

```
    }
]
```

Example — Reconfigurar y cambiar el tamaño de un grupo de instancias en una sola solicitud

En el siguiente ejemplo de JSON se demuestra cómo reconfigurar y cambiar el tamaño de un grupo de instancias en la misma solicitud.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "InstanceCount":5,
      "EC2InstanceIdsToTerminate":["i-123"],
      "ForceShutdown":true,
      "ShrinkPolicy":{
         "DecommissionTimeout":10,
         "InstanceResizePolicy":{
            "InstancesToTerminate":["i-123"],
            "InstancesToProtect":["i-345"],
            "InstanceTerminationTimeout":20
         }
      },
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Configurations":[],
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            }
         }
      ]
   }
]
```

Reconfiguración de un grupo de instancias con el SDK para Java

> **Note**
>
> En los siguientes ejemplos, *<j-2AL4XXXXXX5T9>* sustituya por su ID de clúster y *<ig-1xxxxxxx9>* sustituya por su ID del grupo de instancias.

El siguiente fragmento de código proporciona una nueva configuración para un grupo de instancias mediante AWS SDK para Java.

```
AWSCredentials credentials = new BasicAWSCredentials("access-key", "secret-key");
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);

Map<String,String> hiveProperties = new HashMap<String,String>();
hiveProperties.put("hive.join.emit.interval","1000");
hiveProperties.put("hive.merge.mapfiles","true");

Configuration configuration = new Configuration()
    .withClassification("hive-site")
    .withProperties(hiveProperties);

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withReconfigurationType("MERGE");
    .withConfigurations(configuration);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

Con el siguiente fragmento de código, se elimina una configuración especificada previamente para un grupo de instancias; para ello, se proporciona una matriz vacía de configuraciones.

```
List<Configuration> configurations = new ArrayList<Configuration>();

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
```

```
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withConfigurations(configurations);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

Solución de problemas de reconfiguración de grupos de instancias

Si se produce un error en el proceso de reconfiguración de un grupo de instancias, Amazon EMR revierte la reconfiguración y registra un mensaje de error mediante un evento de Amazon. CloudWatch El evento proporciona un breve resumen del error de reconfiguración. Enumera las instancias en las que la reconfiguración ha fallado y los mensajes de error correspondientes. A continuación se muestra el ejemplo de un mensaje de error.

```
The reconfiguration operation for instance group ig-1xxxxxxx9 in Amazon EMR
 cluster j-2AL4XXXXXX5T9 (ExampleClusterName)
failed at 2021-01-01 00:00 UTC and took 2 minutes to fail. Failed configuration version
 is example12345.
Failure message: Instance i-xxxxxxx1, i-xxxxxxx2, i-xxxxxxx3 failed with message "This
 is an example failure message".
```

Para recopilar más datos sobre un error de reconfiguración, puede consultar los registros de aprovisionamiento de nodos. Esto es especialmente útil cuando recibe un mensaje como el siguiente.

```
i-xxxxxxx1 failed with message “Unable to complete transaction and some changes were
 applied.”
```

On the node

Para acceder a los registros de aprovisionamiento de nodos al conectarse a un nodo

1. Use SSH para conectarse al nodo en el que no se pudo realizar la reconfiguración. Para obtener instrucciones, consulta Conéctate a tu instancia de Linux en la Guía del EC2 usuario de Amazon para instancias de Linux.
2. Navegue hasta el siguiente directorio, que contiene los archivos de registro de aprovisionamiento del nodo.

```
/mnt/var/log/provision-node/
```

3. Abra el subdirectorio `reports` y busque el informe de aprovisionamiento de nodos para la reconfiguración. El `reports` directorio organiza los registros por número de versión de reconfiguración, identificador único universal (UUUUUUUUUID), dirección IP de la EC2 instancia de Amazon y marca de tiempo. Cada informe es un archivo YAML comprimido que contiene información detallada sobre el proceso de reconfiguración.

   A continuación, se muestra un ejemplo del nombre y la ruta de un archivo de informe.

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
xxx.ec2.internal/202104061715.yaml.gz
```

4. Puede examinar un informe con un visor de archivos como `zless`, como se muestra en el siguiente ejemplo.

```
zless 202104061715.yaml.gz
```

Amazon S3

Para acceder a los registros de aprovisionamiento de nodos mediante Amazon S3

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon S3 de https://console.aws.amazon.com/s3/.
2. Abra el bucket de Amazon S3 que ha especificado al configurar el clúster para archivar los archivos de registro.
3. Navegue hasta la siguiente carpeta, que contiene los archivos de registro de aprovisionamiento del nodo:

```
amzn-s3-demo-bucket/elasticmapreduce/<cluster id>/node/<instance id>/provision-
node/
```

4. Abra la carpeta `reports` y busque el informe de aprovisionamiento de nodos para la reconfiguración. La `reports` carpeta organiza los registros por número de versión de reconfiguración, identificador único universal (UUUUUUUUUID), dirección IP de la EC2 instancia de Amazon y marca de tiempo. Cada informe es un archivo YAML comprimido que contiene información detallada sobre el proceso de reconfiguración.

A continuación, se muestra un ejemplo del nombre y la ruta de un archivo de informe.

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
xxx.ec2.internal/202104061715.yaml.gz
```

5. Para ver un archivo de registro, puede descargarlo de Amazon S3 en su máquina local como un archivo de texto. Para obtener instrucciones, consulte Descarga de un objeto.

Cada archivo de registro contiene un informe de aprovisionamiento detallado para la reconfiguración asociada. Para encontrar información sobre los mensajes de error, puede buscar el nivel de registro `err` de un informe. El formato del informe depende de la versión de Amazon EMR del clúster.

En el siguiente ejemplo se muestra información de error para las versiones de Amazon EMR anteriores a la 5.32.0 y la 6.2.0.

```
- !ruby/object:Puppet::Util::Log
      level: !ruby/sym err
      tags:
        - err
      message: "Example detailed error message."
      source: Puppet
      time: 2021-01-01 00:00:00.000000 +00:00
```

Las versiones 5.32.0, 6.2.0 y posteriores de Amazon EMR utilizan el siguiente formato en su lugar.

```
- level: err
  message: 'Example detailed error message.'
  source: Puppet
  tags:
  - err
  time: '2021-01-01 00:00:00.000000 +00:00'
  file:
  line:
```

Almacenamiento de datos de configuración confidenciales en AWS Secrets Manager

Las operaciones de la API de descripción y lista de Amazon EMR que emiten datos de configuración personalizados (como `DescribeCluster` y`ListInstanceGroups`) lo hacen en texto sin formato. Amazon EMR se integra con AWS Secrets Manager para que pueda almacenar sus datos en Secrets Manager y utilizar el ARN del secreto en sus configuraciones. De esta forma, no se transfieren datos de configuración confidenciales a Amazon EMR en texto plano ni se exponen a fuentes externas. APIs Si indica que un par de clave-valor contiene un ARN para un secreto almacenado en Secrets Manager, Amazon EMR recupera este secreto cuando envía los datos de configuración al clúster. Amazon EMR no envía la anotación cuando utiliza externo APIs para mostrar la configuración.

Creación de un secreto

Para crear un secreto, siga los pasos de Creación de un secreto de AWS Secrets Manager en Guía del usuario de AWS Secrets Manager . En el Paso 3, debe elegir el campo Texto no cifrado para ingresar su valor confidencial.

Tenga en cuenta que, si bien Secrets Manager permite que un secreto contenga hasta 65536 bytes, Amazon EMR limita la longitud combinada de la clave de propiedad (excluida la anotación) y el valor secreto recuperado a 1024 caracteres.

Concesión de acceso a Amazon EMR para recuperar el secreto

Amazon EMR utiliza un rol de servicio de IAM para aprovisionar y gestionar los clústeres por usted. El rol de servicio para Amazon EMR define las acciones permitidas para Amazon EMR al aprovisionar recursos y realizar tareas de nivel de servicio que no se llevan a cabo en el contexto de una instancia de Amazon que se ejecuta dentro de un clúster. EC2 Para obtener más información sobre los roles de servicio, consulte Rol de servicio para Amazon EMR (rol de EMR) y Personalización de roles de IAM.

Para permitir que Amazon EMR recupere el valor secreto de Secrets Manager, agregue la siguiente declaración de política a su función de Amazon EMR al lanzar el clúster.

```
{
   "Sid":"AllowSecretsRetrieval",
   "Effect":"Allow",
   "Action":"secretsmanager:GetSecretValue",
   "Resource":[
```

```
      "arn:aws:secretsmanager:<region>:<aws-account-id>:secret:<secret-name>"
   ]
}
```

Si crea el secreto con una administrada por el cliente AWS KMS key, también debe agregar `kms:Decrypt` permiso al rol de Amazon EMR para la clave que utilice. Para obtener más información, consulte Autenticación y control de acceso para AWS Secrets Manager en la Guía del usuario de AWS Secrets Manager .

Uso del secreto en una clasificación de configuración

Puede agregar la anotación `EMR.secret@` a cualquier propiedad de configuración para indicar que su par de clave-valor contiene un ARN de un secreto almacenado en Secrets Manager.

En el siguiente ejemplo se muestra cómo proporcionar el ARN de un secreto en una clasificación de configuración:

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "EMR.secret@presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
   }
}
```

Al crear el clúster y enviar la configuración anotada, Amazon EMR valida las propiedades de la configuración. Si su configuración es válida, Amazon EMR elimina la anotación de la configuración y recupera el secreto de Secrets Manager para crear la configuración real antes de aplicarla al clúster:

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "presto.s3.secret-key":"<my-secret-key-retrieved-from-Secrets-Manager>"
   }
}
```

Al llamar a una acción como `DescribeCluster`, Amazon EMR devuelve la configuración de la aplicación actual en el clúster. Si una propiedad de configuración de la aplicación está marcada

como que contiene el ARN de un secreto, la configuración de la aplicación devuelta por la llamada a `DescribeCluster` contiene el ARN y no el valor del secreto. Esto garantiza que el valor del secreto solo esté visible en el clúster:

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
   }
}
```

Actualización del valor del secreto

Amazon EMR recupera el valor del secreto de una configuración anotada cada vez que el grupo de instancias adjunto se inicia, se reconfigura o cambia de tamaño. Puede usar Secrets Manager para modificar el valor de un secreto utilizado en la configuración de un clúster en ejecución. Cuando lo haga, puede enviar una solicitud de reconfiguración a cada grupo de instancias en el que desee recibir el valor actualizado. Para obtener más información sobre cómo reconfigurar un grupo de instancias y los aspectos que debe tener en cuenta al hacerlo, consulte Reconfiguración de un grupo de instancias en un clúster en ejecución.

Configuración de aplicaciones para utilizar una máquina virtual de Java específica

Las versiones de Amazon EMR tienen diferentes versiones predeterminadas de la máquina virtual Java (JVM). En esta página se explica la compatibilidad de JVM con diferentes versiones y aplicaciones.

Consideraciones

Para obtener información sobre las versiones compatibles de Java para aplicaciones, consulte las páginas de las aplicaciones en la Guía de la versiones de Amazon EMR.

- Amazon EMR solo admite la ejecución de una versión de tiempo de ejecución en un clúster y no admite la ejecución de diferentes nodos o aplicaciones en diferentes versiones de tiempo de ejecución en el mismo clúster.

- En el caso de Amazon EMR 7.x, la máquina virtual Java (JVM) predeterminada es Java 17 para las aplicaciones compatibles con Java 17, con la excepción de Apache Livy. Para obtener información sobre las versiones compatibles del JDK para aplicaciones, consulte la Guía de versiones de Amazon EMR.
- A partir de Amazon EMR 7.1.0, Flink admite y está configurado en Java 17 de forma predeterminada. Para usar una versión diferente del tiempo de ejecución de Java, anule la configuración de `flink-conf`. Para obtener más información sobre cómo configurar Flink para que utilice Java 8 o Java 11, consulte Configurar Flink para que se ejecute con Java 11.
- En el caso de las versiones 5.x y 6.x de Amazon EMR, la máquina virtual Java (JVM) predeterminada es Java 8.
  - Para las versiones 6.12.0 y posteriores de Amazon EMR, algunas aplicaciones también son compatibles con Java 11 y 17.
  - En el caso de las versiones 6.9.0 y posteriores de Amazon EMR, Trino es compatible con Java 17 de forma predeterminada. Para obtener más información sobre Java 17 con Trino, consulte Trino updates to Java 17 en el blog de Trino.

Tenga en cuenta las siguientes consideraciones específicas de la aplicación al elegir la versión del tiempo de ejecución:

Notas de configuración de Java específicas de aplicaciones

Aplicación	Notas de configuración de Java
Spark	Para ejecutar Spark con una versión de Java no predeterminada, debe configurar tanto Spark como Hadoop. Para ver ejemplos, consulta Anulación de la JVM. • Configure `JAVA_HOME` en `spark-env` para actualizar el tiempo de ejecución de Java de los procesos de la instancia principal . Por ejemplo, spark-submit, spark-shell y el servidor de historial de Spark. • Modifique la configuración de Hadoop para actualizar el tiempo de ejecución de Java de

Aplicación	Notas de configuración de Java
	los ejecutores de Spark y YARN `ApplicationMaster` .
Spark RAPIDS	Puede ejecutar RAPIDS con la versión de Java configurada para Spark.
Iceberg	Puede ejecutar Iceberg con la versión de Java configurada de la aplicación que lo utiliza.
Delta	Puede ejecutar Delta con la versión Java configurada de la aplicación que la utiliza.
Hudi	Puede ejecutar Hudi con la versión Java configurada de la aplicación que la utiliza.
Hadoop	Para actualizar la JVM para Hadoop, modifique `hadoop-env` . Para ver ejemplos, consulta Anulación de la JVM.
Hive	Para establecer la versión de Java en 11 o 17 para Hive, establezca la configuración de la JVM de Hadoop en la versión de Java que desee usar.
HBase	Para actualizar la JVM, modifique. HBase `hbase-env` De forma predeterminada, Amazon EMR establece la HBase JVM en función de la configuración de la JVM para Hadoop, a menos que anule la configuración de. `hbase-env` Para ver ejemplos, consulta Anulación de la JVM.

Aplicación	Notas de configuración de Java
Flink	Para actualizar la JVM para Flink, modifique `flink-conf` . De forma predeterminada, Amazon EMR establece la JVM para Flink en función de la configuración de la JVM para Hadoop, a menos que anule la configuración de `flink-conf` . Para obtener más informaci ón, consulte Configuración de Flink para que se ejecute con Java 11.
Oozie	A fin de configurar Oozie para que se ejecute en Java 11 o 17, configure el servidor de Oozie, Oozie LauncherAM Launcher AM y cambie las configuraciones de los trabajos y los ejecutables del cliente. También puede configurar `EmbeddedOozieServer` para que se ejecute en Java 17. Para obtener más información, consulte Configurar la versión Java para Oozie.
Pig	Pig solo es compatible con Java 8. No puede usar Java 11 ni 17 con Hadoop y ejecutar Pig en el mismo clúster.

Anulación de la JVM

A fin de anular la configuración de la JVM para una versión de Amazon EMR (por ejemplo, para usar Java 17 con un clúster que utilice la versión 6.12.0 de Amazon EMR), ingrese la configuración `JAVA_HOME` en su clasificación de entorno, que es *`application`*`-env` para todas las aplicaciones excepto Flink. En el caso de Flink, la clasificación de entorno es `flink-conf`. Para ver los pasos de configuración del tiempo de ejecución de Java con Flink, consulte Configuración de Flink para que se ejecute con Java 11.

Temas

- Anulación de la configuración de la JVM con Apache Spark

- Anulación de la configuración de JVM con Apache HBase
- Anulación de la configuración de la JVM con Apache Hadoop y Hive

Anulación de la configuración de la JVM con Apache Spark

Cuando usa Spark con las versiones 6.12 y posteriores de Amazon EMR, puede configurar el entorno para que los ejecutores usen Java 11 o 17. Y cuando utiliza Spark con versiones de Amazon EMR anteriores a 5.x y escribe un controlador para enviarlo en modo clúster, el controlador utiliza Java 7. Sin embargo, defina el entorno para asegurarse de que los ejecutores utilizan Java 8.

Para sustituir la JVM por Spark, debe establecer las clasificaciones de Hadoop y Spark.

```
{
"Classification": "hadoop-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
            }
        ],
        "Properties": {}
    },
    {
"Classification": "spark-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
            }
        ],
        "Properties": {}
    }
```

El siguiente ejemplo muestra cómo añadir los parámetros de configuración necesarios para EMR 7.0.0+ a fin de garantizar un uso uniforme de la versión Java en todos los componentes.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.executorEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0",
      "spark.yarn.appMasterEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
    }
  },
  {
    "Classification": "hadoop-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
  },
  {
    "Classification": "spark-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
  }
]
```

Anulación de la configuración de JVM con Apache HBase

A fin de configurar HBase para que utilice Java 11, puede establecer la siguiente configuración al lanzar el clúster.

```
[
```

```
  {
    "Classification": "hbase-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/jre-11"
        }
      }
    ],
    "Properties": {}
  }
]
```

Anulación de la configuración de la JVM con Apache Hadoop y Hive

En el siguiente ejemplo se muestra cómo configurar la JVM en la versión 17 para Hadoop y Hive.

```
[
    {
        "Classification": "hadoop-env",
            "Configurations": [
                {
                    "Classification": "export",
                    "Configurations": [],
                    "Properties": {
                        "JAVA_HOME": "/usr/lib/jvm/jre-17"
                    }
                }
        ],
        "Properties": {}
    }
]
```

Puertos de servicio

A continuación se indican los puertos de servicio de YARN y HDFS. Estos ajustes reflejan los valores predeterminados de Hadoop. Otros servicios de aplicaciones se alojan en puertos predeterminados a menos que se documente lo contrario. Para obtener más información, consulte la documentación del proyecto de la aplicación.

Configuración de puertos para YARN y HDFS

Opción	Nombre de host/puerto
fs.default.name	predeterminado (hdfs://*emrDeterminedIP* :8020)
dfs.datanode.address	predeterminado (0.0.0.0:50010)
dfs.datanode.http.address	predeterminado (0.0.0.0:50075)
dfs.datanode.https.address	predeterminado (0.0.0.0:50475)
dfs.datanode.ipc.address	predeterminado (0.0.0.0:50020)
dfs.http.address	predeterminado (0.0.0.0:50070)
dfs.https.address	predeterminado (0.0.0.0:50470)
dfs.secondary.http.address	predeterminado (0.0.0.0:50090)
yarn.nodemanager.address	predeterminado (${yarn.nodemanager.hostname}:0)
yarn.nodemanager.localizer.address	predeterminado (${yarn.nodemanager.hostname}:8040)
yarn.nodemanager.webapp.address	predeterminado (${yarn.nodemanager.hostname}:8042)
yarn.resourcemanager.address	predeterminado (${yarn.resourcemanager.hostname}:8032)
yarn.resourcemanager.admin.address	predeterminado (${yarn.resourcemanager.hostname}:8033)
yarn.resourcemanager.resource-tracker.address	predeterminado (${yarn.resourcemanager.hostname}:8031)
yarn.resourcemanager.scheduler.address	predeterminado (${yarn.resourcemanager.hostname}:8030)

Opción	Nombre de host/puerto
`yarn.resourcemanager.webapp.address`	predeterminado (`${yarn.resourceman ager.hostname}:8088 `)
`yarn.web-proxy.address`	predeterminado (`no-value`)
`yarn.resourcemanager.hostname`	*emrDeterminedIP*

> **Note**
>
> El término *emrDeterminedIP* es una dirección IP generada por el plano de control de Amazon EMR. En la versión más reciente, esta convención se ha eliminado, excepto para la configuración `yarn.resourcemanager.hostname` y `fs.default.name`.

Usuarios de la aplicación

Las aplicaciones ejecutarán procesos como su propio usuario. Por ejemplo, Hive JVMs se ejecuta como usuario`hive`, MapReduce JVMs se ejecuta como`mapred`, etc. Esto se demuestra en el siguiente ejemplo del estado de proceso.

```
USER       PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
hive      6452  0.2  0.7 853684 218520 ?       Sl   16:32   0:13 /usr/lib/jvm/
java-openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-
metastore.log -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop
hive      6557  0.2  0.6 849508 202396 ?       Sl   16:32   0:09 /usr/lib/jvm/java-
openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-server2.log
 -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop/l
hbase     6716  0.1  1.0 1755516 336600 ?      Sl   Jun21   2:20 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_master -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hbase     6871  0.0  0.7 1672196 237648 ?      Sl   Jun21   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_thrift -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hdfs      7491  0.4  1.0 1719476 309820 ?      Sl   16:32   0:22 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_namenode -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-hdfs -
Dhadoop.log.file=hadoop-hdfs-namenode-ip-10-71-203-213.log -Dhadoo
```

```
yarn      8524  0.1  0.6 1626164 211300 ?      Sl   16:33   0:05 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_proxyserver -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn -
Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-yarn-
yarn      8646  1.0  1.2 1876916 385308 ?      Sl   16:33   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_resourcemanager -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn
 -Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-y
mapred    9265  0.2  0.8 1666628 260484 ?      Sl   16:33   0:12 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_historyserver -Xmx1000m -Dhadoop.log.dir=/usr/lib/hadoop/logs -
Dhadoop.log.file=hadoop.log -Dhadoop.home.dir=/usr/lib/hadoop
```

Comprobación de dependencias mediante el repositorio de artefactos de Amazon EMR

Puede utilizar el repositorio de artefactos de Amazon EMR para compilar el código de trabajos de Apache Hive y Apache Hadoop con las versiones exactas de las bibliotecas y dependencias que están disponibles con determinadas versiones de Amazon EMR, a partir de la versión 5.18.0 de Amazon EMR. La compilación con los artefactos de Amazon EMR del repositorio ayuda a evitar problemas de rutas de clases en tiempo de ejecución al asegurarse de que las versiones de las bibliotecas con las que se ha compilado el trabajo son exactamente las mismas versiones proporcionadas en el clúster durante el tiempo de ejecución. En la actualidad, los artefactos de Amazon EMR solo están disponibles para las compilaciones de Maven.

Para obtener acceso al repositorio de artefactos, añada la URL del repositorio al archivo de configuración de Maven o al archivo de configuración `pom.xml` de un proyecto específico. A continuación, puede especificar las dependencias en la configuración del proyecto. Para las versiones de las dependencias, utilice la versión que se indica en la sección Versiones de los componentes para la versión que desee en Versiones de lanzamiento de Amazon EMR 5.x. Por ejemplo, las versiones de los componentes para la versión más reciente de Amazon EMR están disponibles en the section called “Versiones de los componentes de la versión 5.36.2”. Si un artefacto de un proyecto no aparece en la lista de la sección Versiones de los componentes, especifique la versión que se muestra para Hive y Hadoop en esa versión. Por ejemplo, en el caso de los componentes de Hadoop en Amazon EMR versión 5.18.0, la versión es `2.8.4-amzn-1`.

La URL del repositorio de artefactos tiene la siguiente sintaxis:

```
https://s3-endpoint/region-ID-emr-artifacts/emr-release-label/repos/maven/
```

- *`s3-endpoint`* es el punto final de Amazon Simple Storage Service (Amazon S3) de la región del repositorio *`region-ID`* y es la región correspondiente. Por ejemplo, `s3.us-west-1.amazonaws.com` y `us-west-1`. Para obtener más información, consulte Puntos de conexión para Amazon S3 en Referencia general de Amazon Web Services. No hay diferencia en los artefactos entre regiones, por lo que puede especificar la región que le resulte más cómoda para su entorno de desarrollo.
- *`emr-release-label`* es la etiqueta de lanzamiento del clúster Amazon EMR que ejecutará el código. Las etiquetas de versión tienen el formato `emr-`*`x.x.x`*, como `emr-5.36.2`. Una serie de versiones de EMR puede incluir varias versiones. Por ejemplo, si utiliza la versión 5.24.1 de EMR,

utilice la primera etiqueta de versión de EMR dentro de la serie 5.24 `emr-5.24.0`, en la URL del repositorio de artefactos:

```
https://s3-endpoint/region-ID-emr-artifacts/emr-5.24.0/repos/maven/
```

Example Configuración de pom.xml para Maven

El siguiente archivo pom.xml de ejemplo configura un proyecto Maven para que se compile con los artefactos de Apache Hadoop y Apache Hive de emr-5.18.0, utilizando el repositorio de artefactos de us-west-1. Las versiones de las instantáneas no están disponibles en el repositorio de artefactos, por lo que las instantáneas están deshabilitadas en el archivo pom.xml. Los puntos suspensivos (`...`) del ejemplo siguiente indican la omisión de otros parámetros de configuración. No los copie en su proyecto de Maven.

```
<project>
 ...
 <repositories>
  ...
  <repository>
   <id>emr-5.18.0-artifacts</id>
   <name>EMR 5.18.0 Releases Repository</name>
   <releases>
    <enabled>true</enabled>
   </releases>
   <snapshots>
    <enabled>false</enabled>
   </snapshots>
   <url>https://s3.us-west-1.amazonaws.com/us-west-1-emr-artifacts/emr-5.18.0/repos/
maven/</url>
  </repository>
  ...
 </repositories>
 ...
 <dependencies>
  ...
  <dependency>
   <groupId>org.apache.hive</groupId>
   <artifactId>hive-exec</artifactId>
   <version>2.3.3-amzn-2</version>
  </dependency>
  <dependency>
```

```
    <groupId>org.apache.hadoop</groupId>
    <artifactId>hadoop-common</artifactId>
    <version>2.8.4-amzn-1</version>
   </dependency>
   ...
  </dependencies>

</project>
```

Sistema de archivos de EMR (EMRFS)

El sistema de archivos de EMR (EMRFS) es una implementación de HDFS que utilizan todos los clústeres de Amazon EMR para leer y escribir archivos normales de Amazon EMR directamente en Amazon S3. EMRFS proporciona la comodidad de almacenar datos persistentes en Amazon S3 para su uso con Hadoop, al tiempo que proporciona características como el cifrado de datos.

El cifrado de datos permite cifrar los objetos que EMRFS escribe en Amazon S3 y permite a EMRFS trabajar con los objetos cifrados en Amazon S3. Si utiliza la versión 4.8.0 o posterior de Amazon EMR, puede utilizar los ajustes de seguridad para configurar el cifrado para los objetos de EMRFS en Amazon S3, junto con otros ajustes de cifrado. Para obtener más información, consulte Opciones de cifrado. Si utiliza una versión anterior de Amazon EMR, puede configurar manualmente los ajustes de cifrado. Para obtener más información, consulte Especificación del cifrado de Amazon S3 mediante propiedades de EMRFS.

Amazon S3 ofrece una coherencia sólida de lectura después de escritura para todas las operaciones GET, PUT y LIST en todas las Regiones de AWS. Esto significa que lo que escriba con EMRFS es lo que leerá en Amazon S3, sin que ello afecte al rendimiento. Para obtener más información, consulte el Modelo de consistencia de datos de Amazon S3.

Si utiliza la versión 5.10.0 o posterior de Amazon EMR, puede utilizar roles de IAM diferentes para las solicitudes de EMRFS a Amazon S3 en función de los usuarios o los grupos del clúster, o de la ubicación de los datos de EMRFS en Amazon S3. Para obtener más información, consulte Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3.

> **Warning**
>
> Antes de activar la ejecución especulativa para los clústeres de Amazon EMR que ejecutan trabajos de Apache Spark, consulte la siguiente información.
> EMRFS incluye el confirmador optimizado de EMRFS para S3 es una OutputCommitter implementación que está optimizada para la escritura de archivos en Amazon S3 cuando se utiliza EMRFS. Si activa la característica de ejecución especulativa de Apache Spark con aplicaciones que escriben datos en Amazon S3 y no utilizan el confirmador optimizado para S3 de EMRFS, es posible que se produzcan los problemas de corrección de datos descritos en SPARK-10063. Esto puede ocurrir si utiliza versiones de Amazon EMR anteriores a la versión 5.19 de Amazon EMR o si escribe archivos en Amazon S3 con formatos como ORC y CSV. Estos formatos no son compatibles con el confirmador optimizado para S3 de EMRFS.

Para obtener una lista completa de los requisitos para usar el confirmador optimizado para S3 de EMRFS, consulte Requisitos del confirmador optimizado para S3 de EMRFS.
La escritura directa de EMRFS se suele utilizar cuando no se admite el confirmador optimizado para S3 de EMRFS, por ejemplo, cuando se escribe lo siguiente:

- Un formato de salida distinto de Parquet, como ORC o texto.
- Archivos de Hadoop que utilizan la API RDD de Spark.
- Parquet con Hive. SerDe Consulte Hive metastore Parquet table conversion.

La escritura directa de EMRFS no se usa en los siguientes escenarios:

- Cuando no está habilitado el confirmador optimizado para S3 de EMRFS. Consulte Requisitos para el confirmador optimizado para S3 de EMRFS.
- Al escribir particiones dinámicas con el partitionOverwriteMode valor de dynamic.
- Al escribir en ubicaciones de particiones personalizadas, como ubicaciones que no se ajustan a la convención de ubicación de particiones predeterminada de Hive.
- Cuando se utilizan sistemas de archivos distintos de EMRFS, como cuando se escribe en HDFS o se utiliza el sistema de archivos S3A.

Para determinar si su aplicación utiliza la escritura directa en Amazon EMR 5.14.0 o una versión posterior, habilite el registro INFO de Spark. Si hay una línea de registro que contiene el texto “Direct Write: ENABLED” en los registros del controlador de Spark o en los registros del contenedor del ejecutor de Spark, su aplicación de Spark ha escrito mediante escritura directa.
De forma predeterminada, la ejecución especulativa está activada `OFF` en Amazon EMRclusters. Le recomendamos encarecidamente que no active la ejecución especulativa si se cumplen las dos condiciones siguientes:

- Está escribiendo datos en Amazon S3.
- Los datos se escriben en un formato distinto de Apache Parquet o en el formato Apache Parquet sin utilizar el confirmador optimizado para S3 de EMRFS.

Si activa la ejecución especulativa de Spark y escribe datos en Amazon S3 mediante la escritura directa de EMRFS, es posible que sufra una pérdida de datos intermitente. Cuando

escribe datos en HDFS o escribe datos en Parquet mediante el confirmador optimizado para S3 de EMRFS, Amazon EMR no utiliza la escritura directa y este problema no ocurre.
Si necesita escribir datos en formatos que utilizan la escritura directa de EMRFS desde Spark a Amazon S3 y utilizar la ejecución especulativa, le recomendamos que escriba en HDFS y, a continuación, transfiera los archivos de salida a Amazon S3 mediante S3DistCp.

Temas

- Consistent view
- Autorización de acceso a los datos de EMRFS en Amazon S3
- Administrar el AWS Security Token Service punto de conexión de predeterminado
- Especificación del cifrado de Amazon S3 mediante propiedades de EMRFS

Consistent view

Warning

El 1 de junio de 2023 será la fecha de fin de soporte estándar para la vista coherente de EMRFS en futuras versiones de Amazon EMR. La vista coherente de EMRFS seguirá funcionando en las versiones existentes.

Con el lanzamiento de la read-after-write coherencia sólida de Amazon S3 el 1 de diciembre de 2020, ya no necesita usar la vista coherente de EMRFS (CV de EMRFS) con sus clústeres de Amazon EMR. La CV de EMRFS es una característica opcional que permite a los clústeres de Amazon EMR comprobar la lista y la consistencia de los objetos de read-after-write Amazon S3. Cuando crea un clúster y la CV de EMRFS está activada, Amazon EMR crea una base de datos de Amazon DynamoDB para almacenar los metadatos de los objetos que utiliza para realizar un seguimiento de la coherencia de las listas y de los objetos de S3. read-after-write Ahora puede desactivar la CV de EMRFS y eliminar la base de datos de DynamoDB que utiliza para no acumular costos adicionales. Los siguientes procedimientos explican cómo comprobar la característica de CV, desactivarla y eliminar la base de datos de DynamoDB que utiliza la característica.

Para comprobar si está utilizando la característica de CV de EMRFS

1. Navegue hasta la pestaña Configuración. Si el clúster tiene la siguiente configuración, utiliza CV de EMRFS.

   ```
   Classification=emrfs-site,Property=fs.s3.consistent,Value=true
   ```

2. Como alternativa, puede utilizarla AWS CLI para describir el clúster con la API. `describe-cluster` Si el resultado contiene `fs.s3.consistent: true`, el clúster usa la CV de EMRFS.

Para desactivar CV de EMRFS en sus clústeres de Amazon EMR

Para desactivar la característica de CV de EMRFS, utilice una de las tres opciones siguientes. Debe probar estas opciones en su entorno de pruebas antes de aplicarlas a sus entornos de producción.

1. Para detener su clúster existente e iniciar uno nuevo sin las opciones de la CV de EMRFS

   a. Antes de detener el clúster, asegúrese de hacer una copia de seguridad de los datos y de avisar a los usuarios.

   b. Para detener el clúster, siga las instrucciones de la sección Terminar un clúster.

   c. Si utiliza la consola de Amazon EMR para crear un clúster nuevo, vaya a Opciones avanzadas. En la sección Editar configuración de software, anule la selección de la opción para activar la CV de EMRFS. Si la casilla que corresponde a la vista coherente de EMRFS está disponible, manténgala desmarcada.

   d. Si usa AWS CLI para crear un clúster nuevo con la `create-clusterAPI`, no utilice la `--emrfs` opción, que activa la CV de EMRFS.

   e. Si utiliza un SDK o AWS CloudFormation para crear un clúster nuevo, no utilice ninguna de las configuraciones que se indican en Configurar la vista coherente.

2. Para clonar un clúster y eliminar el CV de EMRFS

   a. En la consola de Amazon EMR, elija el clúster que usa CV de EMRFS.

   b. En la parte superior de la página Detalles del clúster, elija Clonar.

   c. Elija Anterior y vaya a Paso 1: software y pasos.

   d. En Editar configuración de software, elimine la CV de EMRFS. En Editar configuración, elimine las siguientes configuraciones de la clasificación `emrfs-site`. Si está cargando JSON desde un bucket de S3, debe modificar su objeto de S3.

```
[
 {"classification":
  "emrfs-site",
  "properties": {
   "fs.s3.consistent.retryPeriodSeconds":"10",
   "fs.s3.consistent":"true",
   "fs.s3.consistent.retryCount":"5",
   "fs.s3.consistent.metadata.tableName":"EmrFSMetadata"
  }
 }
]
```

3. Para eliminar el CV de EMRFS de un clúster que usa grupos de instancias

   a. Utilice el siguiente comando para comprobar si un único clúster de EMR utiliza la tabla de DynamoDB asociada al CV de EMRFS o si varios clústeres comparten la tabla. El nombre de la tabla se especifica en `fs.s3.consistent.metadata.tableName`, tal y como se describe en Configuración de la vista coherente. El nombre de tabla predeterminado que utiliza la CV de EMRFS es `EmrFSMetadata`.

   ```
   aws emr describe-cluster --cluster-id j-XXXXX | grep
    fs.s3.consistent.metadata.tableName
   ```

   b. Si el clúster no comparte la base de datos de DynamoDB con otro clúster, utilice el siguiente comando para volver a configurar el clúster y desactivar la CV de EMRFS. Para obtener más información, consulte la sección Reconfiguración de un grupo de instancias en un clúster en ejecución.

   ```
   aws emr modify-instance-groups --cli-input-json file://disable-emrfs-1.json
   ```

   Este comando abre el archivo que desea modificar. Modifique el archivo con las siguientes configuraciones.

   ```
   {
    "ClusterId": "j-xxxx",
    "InstanceGroups": [
     {
      "InstanceGroupId": "ig-xxxx",
      "Configurations": [
       {
   ```

```
      "Classification": "emrfs-site",
      "Properties": {
       "fs.s3.consistent": "false"
      },
      "Configurations": []
     }
    ]
   }
 ]
}
```

c. Si su clúster comparte la tabla de DynamoDB con otro clúster, desactive la CV de EMRFS en todos los clústeres cuando ningún clúster modifique ningún objeto de la ubicación de S3 compartida.

Para eliminar los recursos de Amazon DynamoDB asociados al CV de EMRFS

Después de eliminar el CV de EMRFS de sus clústeres de Amazon EMR, elimine los recursos de DynamoDB asociados al CV de EMRFS. Hasta que lo haga, seguirá incurriendo en cargos de DynamoDB asociados a la CV de EMRFS.

1. Compruebe las CloudWatch métricas de la tabla de DynamoDB y confirme que ningún clúster la utilice.
2. Elimine la tabla de DynamoDB.

```
aws dynamodb delete-table --table-name <your-table-name>
```

Para eliminar los recursos de Amazon SQS asociados al CV de EMRFS

1. Si configuró el clúster para enviar notificaciones de incoherencia a Amazon SQS, puede eliminar todas las colas de SQS.
2. Encuentre el nombre de cola de Amazon SQS se especifica en `fs.s3.consistent.notification.SQS.queueName`, tal y como se describe en Configuración de la vista coherente. El formato de nombre de cola predeterminado es `EMRFS-Inconsistency-<j-cluster ID>`.

```
aws sqs list-queues | grep 'EMRFS-Inconsistency'
aws sqs delete-queue -queue-url <your-queue-url>
```

Para dejar de usar la CLI de EMRFS

- La CLI de EMRFS administra los metadatos que genera la CV de EMRFS. Cuando el soporte estándar para la CV de EMRFS llegue a su fin en futuras versiones de Amazon EMR, el soporte para la CLI de EMRFS también lo hará.

Temas

- Habilitación de la vista coherente
- Seguimiento de objetos de EMRFS mediante la vista consistente de Amazon S3
- Lógica de reintentos
- Metadatos de la vista consistente de EMRFS
- Configurar notificaciones de consistencia para Amazon CloudWatch SQS y Amazon SQS
- Configuración de la vista coherente
- La Referencia de los comandos de la CLI de EMRFS

Habilitación de la vista coherente

Puede habilitar el cifrado del servidor de Amazon S3 o la vista coherente de EMRFS mediante la AWS Management Console AWS CLI, o la clasificación de configuración. `emrfs-site`

Para configurar la vista consistente utilizando la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. Elija los valores de Step 1: Software and Steps (Paso 1: Software y pasos) y Step 2: Hardware (Paso 2: Hardware).
4. En Step 3: General Cluster Settings (Paso 3: Configuración general del clúster), en Additional Options (Opciones adicionales), elija EMRFS consistent view (Vista consistente de EMRFS).
5. En EMRFS Metadata store (Almacén de metadatos de EMRFS), escriba el nombre del almacén de metadatos. El valor predeterminado es **`EmrFSMetadata`**. Si la FSMetadata tabla Emr no existe, se crea automáticamente en DynamoDB.

> **Note**
>
> Amazon EMR no elimina automáticamente los metadatos de EMRFS desde DynamoDB cuando se termina el clúster.

6. En Number of retries (Número de reintentos), escriba un valor entero. Si se detecta una inconsistencia, EMRFS intenta llamar a Amazon S3 este número de veces. El valor predeterminado es **5**.
7. En Retry period (in seconds) (Período de reintento (en segundos)), escriba un valor entero. Se trata de la cantidad de tiempo que EMRFS espera entre reintentos. El valor predeterminado es **10**.

> **Note**
>
> Los reintentos siguientes utilizan un retardo exponencial.

Para lanzar un clúster con una vista coherente, habilitada mediante la AWS CLI

Le recomendamos que instale la versión actual de AWS CLI. Para descargar la versión más reciente, consulte https://aws.amazon.com/cli/.

- > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --instance-type m5.xlarge --instance-count 3 --emrfs
 Consistent=true \
--release-label emr-7.9.0 --ec2-attributes KeyName=myKey
```

Para comprobar si la vista coherente está habilitada, utilice la AWS Management Console

- Para comprobar si la vista consistente está habilitada en la consola, vaya a Cluster List (Lista de clústeres) y seleccione el nombre del clúster para ver los Cluster Details (Detalles del clúster). El campo "EMRFS consistent view" tiene un valor de `Enabled` o `Disabled`.

Para comprobar si la vista consistente está habilitada examinando el archivo **`emrfs-site.xml`**

- Puede comprobar si la consistencia está habilitada inspeccionando el archivo de configuración `emrfs-site.xml` en el nodo principal del clúster. Si el valor booleano para `fs.s3.consistent` se define en `true` entonces la vista consistente está habilitada para las operaciones del sistema de archivos que implican Amazon S3.

Seguimiento de objetos de EMRFS mediante la vista consistente de Amazon S3

EMRFS crea una vista consistente de los objetos en Amazon S3 agregando información sobre estos objetos a los metadatos de EMRFS. EMRFS añade estas listas a sus metadatos cuando:

- Un objeto escrito por EMRFS durante el transcurso de un trabajo de Amazon EMR.
- Un objeto se sincroniza con o se importa a los metadatos de EMRFS utilizando la CLI de EMRFS.

Los objetos leídos por EMRFS no se añaden automáticamente a los metadatos. Cuando EMRFS elimina un objeto, su entrada permanece en los metadatos con el estado eliminado hasta que dicha entrada se purga utilizando la CLI de EMRFS. Para obtener más información sobre la CLI, consulte La Referencia de los comandos de la CLI de EMRFS. Para obtener más información sobre la eliminación de listados en los metadatos de EMRFS, consulte Metadatos de la vista consistente de EMRFS.

Por cada operación de Amazon S3, EMRFS comprueba los metadatos para obtener información sobre el conjunto de objetos en la vista consistente. Si EMRFS detecta que Amazon S3 es inconsistente durante una de estas operaciones, vuelve a intentar la operación, de acuerdo con los parámetros definidos en las propiedades de configuración de `emrfs-site`. Cuando EMRFS agota los reintentos, genera una excepción `ConsistencyException` o registra la excepción y continúa el flujo de trabajo. Para obtener más información acerca de la lógica de reintentos, consulte Lógica de reintentos. Puede encontrar `ConsistencyExceptions` en sus registros, por ejemplo:

- listStatus: ningún objeto de Amazon S3 para elemento de metadatos `/S3_bucket/dir/object`
- getFileStatus: La clave `dir/file` está presente en los metadatos, pero no en Amazon S3

Si elimina un objeto directamente desde Amazon S3 y se está realizando su seguimiento en la vista consistente de EMRFS, EMRFS trata dicho objeto como inconsistente, ya que sigue apareciendo en los metadatos como presente en Amazon S3. Si los metadatos dejan de estar sincronizados con los objetos cuyo seguimiento realiza EMRFS en Amazon S3, puede utilizar el subcomando sync de la CLI de EMRFS para restablecer el estado de los metadatos de Amazon S3. Para descubrir discrepancias entre los metadatos y Amazon S3, utilice el comando diff. Por último, EMRFS solo tiene una vista consistente de los objetos a los que se hace referencia en los metadatos; puede haber otros objetos en la misma ruta de Amazon S3 de los que no se está realizando seguimiento. Cuando EMRFS muestra un listado de los objetos de una ruta de Amazon S3, devuelve el superconjunto de los objetos a los que se hace seguimiento en los metadatos y de los que están en esa ruta de Amazon S3.

Lógica de reintentos

EMRFS intenta comprobar la consistencia de la lista de objetos a los que se hace seguimiento en sus metadatos un número de reintentos específico. El valor predeterminado es 5. En caso de que el número de reintentos supere el trabajo de origen se devuelve un error, a menos que `fs.s3.consistent.throwExceptionOnInconsistency` se haya establecido en `false`, donde solo registrará los objetos marcados como inconsistentes. EMRFS utiliza una política de reintento de retardo exponencial de forma predeterminada, pero también puede establecer a una política fija. Es posible que los usuarios también deseen volver a intentar tras un periodo de tiempo determinado antes de proceder con el resto de su trabajo sin lanzar una excepción. Pueden conseguirlo configurando `fs.s3.consistent.throwExceptionOnInconsistency` en `false`, `fs.s3.consistent.retryPolicyType` en `fixed` y `fs.s3.consistent.retryPeriodSeconds` para el valor deseado. El siguiente ejemplo crea un clúster con la consistencia habilitada, que registra las inconsistencias y establece un intervalo de reintento fijo de 10 segundos:

Example Establecimiento de un periodo de reintento en una cantidad fija

```
aws emr create-cluster --release-label emr-7.9.0 \
--instance-type m5.xlarge --instance-count 1 \
```

```
--emrfs Consistent=true,Args=[fs.s3.consistent.throwExceptionOnInconsistency=false,
 fs.s3.consistent.retryPolicyType=fixed,fs.s3.consistent.retryPeriodSeconds=10] --ec2-
attributes KeyName=myKey
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Para obtener más información, consulte Consistent view.

Las configuraciones de EMRFS para el IMDS obtienen llamadas regionales

EMRFS se basa en el IMDS (servicio de metadatos de instancias) para obtener la región de la instancia y los puntos de conexión de Amazon S3, DynamoDB o. AWS KMS Sin embargo, el IMDS tiene un límite en cuanto al número de solicitudes que puede gestionar y las que superen ese límite fallarán. Este límite del IMDS puede provocar errores en la inicialización de EMRFS y provocar un error en la consulta o el comando. Puede utilizar el siguiente mecanismo de reintentos de retroceso exponencial con asignación al azar y propiedades de configuración de una región alternativa en emrfs-site.xml para abordar el escenario en el que todos los reintentos fallan.

```
<property>
    <name>fs.s3.region.retryCount</name>
    <value>3</value>
    <description>
    Maximum retries that would be attempted to get AWS region.
    </description>
</property>
<property>
    <name>fs.s3.region.retryPeriodSeconds</name>
    <value>3</value>
    <description>
    Base sleep time in second for each get-region retry.
    </description>
</property>
<property>
    <name>fs.s3.region.fallback</name>
    <value>us-east-1</value>
    <description>
```

```
    Fallback to this region after maximum retries for getting AWS region have been
 reached.
    </description>
</property>
```

Metadatos de la vista consistente de EMRFS

La vista consistente de EMRFS realiza un seguimiento de la consistencia mediante una tabla de DynamoDB que permite controlar los objetos de Amazon S3 que se han creado o sincronizado con EMRFS. Los metadatos se utilizan para realizar un seguimiento de todas las operaciones (lectura, escritura, actualización y copia) y no se almacena ningún contenido real. Estos metadatos se utilizan para validar si los objetos o los metadatos recibidos desde Amazon S3 coinciden con lo que se espera. Esta confirmación ofrece a EMRFS la capacidad de comprobar la consistencia de lista y read-after-write la consistencia de las listas de nuevos objetos que EMRFS escribe en Amazon S3 u objetos sincronizados con EMRFS. Varios clústeres pueden compartir los mismos metadatos.

Cómo añadir entradas a metadatos

Puede utilizar los subcomandos `sync` o `import` para agregar entradas a los metadatos. `sync` refleja el estado de los objetos de Amazon S3 de una ruta, mientras que `import` se utiliza estrictamente para añadir entradas nuevas a los metadatos. Para obtener más información, consulte La Referencia de los comandos de la CLI de EMRFS.

Cómo comprobar las diferencias entre metadatos y objetos en Amazon S3

Para comprobar si hay diferencias entre los metadatos y Amazon S3, utilice el subcomando `diff` de la CLI de EMRFS. Para obtener más información, consulte La Referencia de los comandos de la CLI de EMRFS.

Cómo saber si las operaciones de metadatos se están limitando

EMRFS establece límites de capacidad de rendimiento predeterminados en los metadatos para sus operaciones de lectura y escritura en 500 y 100 unidades, respectivamente. Un gran número de objetos o buckets podría provocar que las operaciones superen esta capacidad, a partir de la cual se verán limitadas por DynamoDB. Por ejemplo, una aplicación podría hacer que EMRFS lance una `ProvisionedThroughputExceededException` si realiza una operación que supere estos límites de capacidad. Tras la limitación, la herramienta de CLI de EMRFS vuelve a intentar la escritura en la tabla de DynamoDB mediante retroceso exponencial hasta que finalice la operación o cuando alcance el máximo valor de reintento para escribir objetos desde Amazon EMR a Amazon S3.

Puede configurar sus propios límites de la capacidad de rendimiento. Sin embargo, DynamoDB tiene límites de partición estrictos de 3000 unidades de capacidad de lectura RCUs () y 1000 unidades de capacidad de escritura WCUs () por segundo para las operaciones de lectura y escritura. Para evitar `sync` errores provocados por la limitación, le recomendamos que limite el rendimiento de las operaciones de lectura a menos de 3000 RCUs y el de las operaciones de escritura a menos de 1000. WCUs Para obtener instrucciones sobre cómo configurar límites de capacidad de rendimiento personalizados, consulte Configuración de la vista coherente.

También puede ver CloudWatch las métricas de Amazon para los metadatos de EMRFS en la consola de DynamoDB, donde podrá ver el número de solicitudes de lectura y escritura limitadas. Si tiene un valor distinto de cero para las solicitudes limitadas, la aplicación podría potencialmente beneficiarse del aumento de capacidad de rendimiento asignada para operaciones de lectura o escritura. También puede obtener un beneficio de rendimiento si ve que sus operaciones se aproximan a la capacidad máxima de rendimiento asignada en lecturas o escrituras durante un periodo de tiempo prolongado.

Características de rendimiento para operaciones de EMRFS notables

El valor predeterminado para las operaciones de lectura y escritura es de 400 y 100 unidades de capacidad de rendimiento, respectivamente. Las siguientes características de rendimiento le dan una idea del rendimiento necesario para determinadas operaciones. Estas pruebas se realizaron a través de clúster `m3.large` de un único nodo. Todas las operaciones tenían un único hilo. El rendimiento varia en gran medida en función de las características de la aplicación en particular y es posible que tenga que experimentar para optimizar las operaciones del sistema de archivos.

Operación	Promedio read-per-second	Promedio write-per-second
create (objeto)	26,79	6,70
delete (objeto)	10,79	10,79
delete (directorio que contiene 1000 objetos)	21,79	338,40
getFileStatus(objeto)	34,70	0
getFileStatus(directorio)	19,96	0

Operación	Promedio read-per-second	Promedio write-per-second
listStatus (directorio que contiene un objeto)	43,31	0
listStatus (directorio que contiene 10 objetos)	44,34	0
listStatus (directorio que contiene 100 objetos)	84,44	0
listStatus (directorio que contiene 1000 objetos)	308,81	0
listStatus (directorio que contiene 10 000 objetos)	416,05	0
listStatus (directorio que contiene 100 000 objetos)	823,56	0
listStatus (directorio que contiene un millón de objetos)	882,36	0
mkdir (continuo para 120 segundos)	24,18	4,03
mkdir	12,59	0
rename (objeto)	19,53	4,88
rename (directorio que contiene 1000 objetos)	23,22	339,34

Para enviar un paso que elimina datos antiguos de su almacén de metadatos

Es posible que los usuarios deseen eliminar entradas particulares en los metadatos basados en DynamoDB. Esto puede ayudar a reducir los costos de almacenamiento asociados a la tabla. Los usuarios tienen la posibilidad de eliminar manualmente o mediante programación las entradas

particulares a través del subcomando `delete` de la CLI de EMRFS. No obstante, si elimina las entradas desde los metadatos, EMRFS ya no realiza ninguna comprobación de consistencia.

La eliminación mediante programación tras la finalización de un trabajo puede realizarse enviando un último paso al clúster, que ejecuta un comando de la CLI de EMRFS. Por ejemplo, escriba el siguiente comando para enviar un paso a su clúster para eliminar todas las entradas de más de dos días.

```
aws emr add-steps --cluster-id j-2AL4XXXXXX5T9 --steps Name="emrfsCLI",Jar="command-
runner.jar",Args=["emrfs","delete","--time","2","--time-unit","days"]
{
    "StepIds": [
        "s-B12345678902"
    ]
}
```

Utilice el StepId valor devuelto para comprobar los registros para el resultado de la operación.

Configurar notificaciones de consistencia para Amazon CloudWatch SQS y Amazon SQS

Puede habilitar CloudWatch las métricas y los mensajes de Amazon SQS en EMRFS para solucionar problemas de consistencia final de Amazon S3.

CloudWatch

Cuando están habilitadas CloudWatch las métricas, se envía una métrica denominada Incoherencia cada vez que una llamada a la `FileSystem` API devuelve un error debido a consistencia final de Amazon S3.

Para ver CloudWatch las métricas de problemas de consistencia final de Amazon S3

Para ver la métrica de inconsistencia en la CloudWatch consola, seleccione las métricas de EMRFS y, a continuación, seleccione un par JobFlowId/nombre de métrica. Por ejemplo, `j-162XXXXXXM2CU ListStatus`, `j-162XXXXXXM2CU GetFileStatus` etc.

1. Abra la consola en CloudWatch . https://console.aws.amazon.com/cloudwatch/
2. En el Dashboard (Panel), en la sección Metrics (Métricas), elija EMRFS.
3. En el panel Métricas del flujo de trabajos, seleccione uno o más pares de JobFlowId/Nombre de métrica. Aparece una representación gráfica de las métricas en la ventana siguiente.

Amazon SQS

Cuando están habilitadas las notificaciones de Amazon SQS, se crea una cola de Amazon SQS con el nombre `EMRFS-Inconsistency-<jobFlowId>` cuando se inicializa EMRFS. Los mensajes de Amazon SQS se envían en la cola cuando falla una llamada a la API de `FileSystem` debido a consistencia final de Amazon S3. El mensaje contiene información como la API JobFlowId, una lista de rutas incoherentes, un rastreo de pila, etc. Los mensajes se pueden leer a través de la consola de Amazon SQS o mediante el comando `read-sqs` de EMRFS.

Para administrar mensajes de Amazon SQS para problemas de consistencia final de Amazon S3

Los mensajes de Amazon SQS para problemas de consistencia final de Amazon S3 se pueden leer con la CLI de EMRFS. Para leer mensajes desde una cola de Amazon SQS de EMRFS, escriba el comando `read-sqs` y especifique una ubicación de salida en el sistema de archivos local del nodo maestro para el archivo de salida resultante.

También puede eliminar una cola de Amazon SQS de EMRFS utilizando el comando `delete-sqs`.

1. Para leer los mensajes de una cola de Amazon SQS, escriba el siguiente comando. Sustituya *queuename* por el nombre de la cola de Amazon SQS que ha configurado y sustituya por */path/filename* la ruta al archivo de salida:

   ```
   emrfs read-sqs --queue-name queuename --output-file /path/filename
   ```

   Por ejemplo, para leer y generar mensajes de Amazon SQS desde la cola predeterminada, escriba:

   ```
   emrfs read-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU --output-file /path/filename
   ```

   > **Note**
   >
   > También puede utilizar los accesos directos `-q` y `-o` en lugar de `--queue-name` y `--output-file` respectivamente.

2. Para eliminar una cola de Amazon SQS, escriba el comando siguiente:

   ```
   emrfs delete-sqs --queue-name queuename
   ```

Por ejemplo, para eliminar la cola predeterminada, escriba:

```
emrfs delete-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU
```

> **Note**
>
> También puede utilizar el acceso directo `-q` en lugar de `--queue-name`.

Configuración de la vista coherente

Puede configurar ajustes adicionales para la vista consistente mediante las propiedades de configuración de `emrfs-site`. Por ejemplo, puede elegir otra opción predeterminada de rendimiento de DynamoDB suministrando los siguientes argumentos a la opción `--emrfs` de la CLI, utilizando la clasificación de configuración emrfs-site (solo para Amazon EMR versión 4.x y posteriores), o mediante una acción de arranque para configurar el archivo emrfs-site.xml en el nodo maestro:

Example Cambio de valores de lectura y escritura de metadatos predeterminados en el lanzamiento del clúster

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--emrfs Consistent=true,Args=[fs.s3.consistent.metadata.read.capacity=600,\
fs.s3.consistent.metadata.write.capacity=300] --ec2-attributes KeyName=myKey
```

De forma alternativa, utilice el siguiente archivo de configuración y guárdelo de forma local o en Amazon S3:

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.consistent.metadata.read.capacity": "600",
        "fs.s3.consistent.metadata.write.capacity": "300"
      }
    }
 ]
```

Utilice la configuración que ha creado con la siguiente sintaxis:

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hive \
--instance-type m5.xlarge --instance-count 2 --configurations file://./myConfig.json
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Las siguientes opciones se pueden configurar mediante configuraciones o AWS CLI `--emrfs` argumentos. Para obtener información sobre estos argumentos, consulte Referencia de comandos de AWS CLI.

Propiedades **`emrfs-site.xml`** para vista coherente

Propiedad	Valor predeterminado	Descripción
`fs.s3.consistent`	**`false`**	Cuando se establece en **`true`**, esta propiedad configura EMRFS para utilizar DynamoDB para proporcionar consistencia.
`fs.s3.consistent.retryPolicyType`	**`exponential`**	Esta propiedad identifica la política que utilizar a la hora de volver a intentar los problemas de consistencia. Las opciones incluyen: exponential, fixed o none.
`fs.s3.consistent.retryPeriodSeconds`	**`1`**	Esta propiedad establece la duración que esperar entre la reintentos de consistencia.
`fs.s3.consistent.retryCount`	**`10`**	Esta propiedad establece el número máximo de reintento

Propiedad	Valor predeterminado	Descripción
		s cuando se detecta inconsistencia.
`fs.s3.consistent.throwExceptionOnInconsistency`	**`true`**	Esta propiedad determina si lanzar o registrar una excepción de consistencia. Cuando se establece en **`true`**, se lanza una `ConsistencyException` .
`fs.s3.consistent.metadata.autoCreate`	**`true`**	Cuando se establece en **`true`**, esta propiedad habilita la creación automática de tablas de metadatos.
`fs.s3.consistent.metadata.etag.verification.enabled`	**`true`**	Con Amazon EMR 5.29.0, esta propiedad está habilitada de forma predeterminada. Cuando está habilitada, EMRFS utiliza S3 ETags para verificar que los objetos leídos tienen la última versión disponible. Esta característica resulta útil para casos de read-after-update uso en los que los archivos de S3 se sobrescriben conservando el mismo nombre. Esta capacidad de ETag verificación no funciona actualmente con S3 Select.
`fs.s3.consistent.metadata.tableName`	**`EmrFSMetadata`**	Esta propiedad especifica el nombre de la tabla de metadatos en DynamoDB.

Propiedad	Valor predeterminado	Descripción
`fs.s3.consistent.metadata.read.capacity`	**`500`**	Esta propiedad especifica la capacidad de lectura de DynamoDB para aprovisionar cuando se crea la tabla de metadatos.
`fs.s3.consistent.metadata.write.capacity`	**`100`**	Esta propiedad especifica la capacidad de escritura de DynamoDB para aprovisionar cuando se crea la tabla de metadatos.
`fs.s3.consistent.fastList`	**`true`**	Cuando se establece en **`true`**, esta propiedad utiliza varios subprocesos para mostrar un directorio (cuando sea necesario). La consistencia debe estar habilitada para utilizar esta propiedad.
`fs.s3.consistent.fastList.prefetchMetadata`	**`false`**	Cuando se establece en **`true`**, esta propiedad habilita la captura previa de metadatos para los directorios que contengan más de 20 000 elementos.

Propiedad	Valor predeterminado	Descripción
`fs.s3.consistent.notification.CloudWatch`	**`false`**	Cuando se establece en**`true`**, CloudWatch las métricas están habilitadas para las llamadas a la FileSystem API que devuelven un error debido a problemas de consistencia final de Amazon S3.
`fs.s3.consistent.notification.SQS`	**`false`**	Cuando se establece en **`true`**, las notificaciones de consistencia final se envían a una cola de Amazon SQS.
`fs.s3.consistent.notification.SQS.queueName`	**`EMRFS-Inconsistency-<jobFlowId>`**	Cambiar esta propiedad le permite especificar su propio nombre de cola de SQS para mensajes relativos a problemas de consistencia final de Amazon S3.
`fs.s3.consistent.notification.SQS.customMsg`	**`none`**	Esta propiedad le permite especificar información personalizada incluida en mensajes de SQS relativos a problemas de consistencia final de Amazon S3. Si no se especifica ningún valor para esta propiedad, el campo correspondiente en el mensaje está vacío.

Propiedad	Valor predeterminado	Descripción
`fs.s3.consistent.dynamodb.endpoint`	**`none`**	Esta propiedad le permite especificar un punto de enlace de DynamoDB personalizado para sus metadatos de vista consistente.
`fs.s3.useRequesterPaysHeader`	**`false`**	Si se establece en **`true`**, esta propiedad permite que Amazon S3 envíe solicitudes a buckets con la opción de pagador de solicitudes habilitada.

La Referencia de los comandos de la CLI de EMRFS

La CLI de EMRFS se instala de forma predeterminada en todos los nodos principales del clúster creados con la versión 3.2.1 o posterior de Amazon EMR. Puede utilizar la CLI de EMRFS para administrar los metadatos de la vista consistente.

> **Note**
>
> El emrfs comando solo es compatible con la emulación de terminal VT1 00. Sin embargo, podría funcionar con otros modos de emulador del terminal.

comando de nivel superior emrfs

El comando de nivel superior emrfs admite la siguiente estructura.

```
emrfs [describe-metadata | set-metadata-capacity | delete-metadata | create-metadata |
 \
list-metadata-stores | diff | delete | sync | import ] [options] [arguments]
```

Especifique [opciones], con o sin [argumentos] tal y como se describe en la siguiente tabla. Para [opciones] específicas de subcomandos (`describe-metadata`, `set-metadata-capacity`, etc.), consulte cada subcomando a continuación.

[Opciones] para emrfs

Opción	Descripción	Obligatorio
`-a` *`AWS_ACCESS_KEY_ID`* | `--access-key` *`AWS_ACCESS_KEY_ID`*	La clave de AWS acceso de que se utiliza para escribir objetos en Amazon S3 y para crear o tener acceso a un almacén de metadatos en DynamoDB. De forma predeterminada, *`AWS_ACCESS_KEY_ID`* se establece en la clave de acceso utilizada para crear el clúster.	No
`-s` *`AWS_SECRET_ACCESS_KEY`* | `--secret-key` *`AWS_SECRET_ACCESS_KEY`*	La clave AWS secreta de asociada a la clave de acceso de que se utiliza para escribir objetos en Amazon S3 y para crear o tener acceso a un almacén de metadatos en DynamoDB. De forma predeterminada, *`AWS_SECRET_ACCESS_KEY`* se establece en la clave secreta asociada a la clave de acceso; utilizada para crear el clúster.	No
`-v` | `--verbose`	Hace que la salida sea detallada.	No
`-h` | `--help`	Muestra el mensaje de ayuda para el comando `emrfs` con una declaración de uso.	No

subcomando emrfs describe-metadata

[Opciones] para emrfs describe-metadata

Opción	Descripción	Obligatorio

Opción	Descripción	Obligatorio
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *METADATA_NAME* argumento, el valor predeterminado es. `EmrFSMetadata`	No

Example ejemplo de emrfs describe-metadata

En el siguiente ejemplo se describe la tabla de metadatos predeterminados.

```
$ emrfs describe-metadata
EmrFSMetadata
  read-capacity: 400
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 12
```

subcomando emrfs set-metadata-capacity

[Opciones] para emrfs set-metadata-capacity

Opción	Descripción	Obligatorio
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *METADATA_NAME* argumento, el valor predeterminado es. `EmrFSMetadata`	No
`-r` *READ_CAPACITY* | `--read-capacity` *READ_CAPACITY*	La capacidad de rendimiento de lectura solicitada para la tabla de metadatos. Si no se proporciona el *READ_CAPACITY* argumento, el valor por defecto es`400`.	No
		No

Opción	Descripción	Obligatorio
`-w` *`WRITE_CAPACITY`* | `--write-capacity` *`WRITE_CAPACITY`*	La capacidad de rendimiento de escritura solicitada para la tabla de metadatos. Si no se proporciona el *`WRITE_CAPACITY`* argumento, el valor por defecto es`100`.	

Example ejemplo de emrfs set-metadata-capacity

El siguiente ejemplo establece la capacidad de rendimiento de lectura en `600` y la capacidad de escritura en `150` para una tabla de metadatos denominada `EmrMetadataAlt`.

```
$ emrfs set-metadata-capacity --metadata-name EmrMetadataAlt  --read-capacity 600 --
write-capacity 150
  read-capacity: 400
  write-capacity: 100
  status: UPDATING
  approximate-item-count (6 hour delay): 0
```

subcomando emrfs delete-metadata

[Opciones] para emrfs delete-metadata

Opción	Descripción	Obligatorio
`-m` *`METADATA_NAME`* | `--metadata-name` *`METADATA_NAME`*	*`METADATA_NAME`* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *`METADATA_NAME`* argumento, el valor predeterminado es. `EmrFSMetadata`	No

Example ejemplo de emrfs delete-metadata

En el siguiente ejemplo se elimina la tabla de metadatos predeterminados.

```
$ emrfs delete-metadata
```

subcomando emrfs create-metadata

[Opciones] para emrfs create-metadata

Opción	Descripción	Obligatorio
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *METADATA_NAME* argumento, el valor predeterminado es. `EmrFSMetadata`	No
`-r` *READ_CAPACITY* | `--read-capacity` *READ_CAPACITY*	La capacidad de rendimiento de lectura solicitada para la tabla de metadatos. Si no se proporciona el *READ_CAPACITY* argumento, el valor por defecto es`400`.	No
`-w` *WRITE_CAPACITY* | `--write-capacity` *WRITE_CAPACITY*	La capacidad de rendimiento de escritura solicitada para la tabla de metadatos. Si no se proporciona el *WRITE_CAPACITY* argumento, el valor por defecto es`100`.	No

Example ejemplo de emrfs create-metadata

El siguiente ejemplo crea una tabla de metadatos denominada `EmrFSMetadataAlt`.

```
$ emrfs create-metadata -m EmrFSMetadataAlt
Creating metadata: EmrFSMetadataAlt
EmrFSMetadataAlt
  read-capacity: 400
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 0
```

subcomando emrfs list-metadata-stores

El subcomando emrfs list-metadata-stores no tiene opciones.

Example List-metadata-stores ejemplo

El siguiente ejemplo muestra las tablas de metadatos.

```
$ emrfs list-metadata-stores
  EmrFSMetadata
```

subcomando emrfs diff

[Opciones] para emrfs diff

Opción	Descripción	Obligatorio
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *METADATA_NAME* argumento, el valor predeterminado es. `EmrFSMetadata`	No
s3://s3Path	La ruta del bucket de Amazon S3 que se debe comparar con la tabla de metadatos. Los buckets se sincronizan de forma recursiva.	Sí

Example ejemplo de emrfs diff

En el siguiente ejemplo se compara la tabla de metadatos predeterminados con un bucket de Amazon S3.

```
$ emrfs diff s3://elasticmapreduce/samples/cloudfront
BOTH | MANIFEST ONLY | S3 ONLY
DIR elasticmapreduce/samples/cloudfront
DIR elasticmapreduce/samples/cloudfront/code/
DIR elasticmapreduce/samples/cloudfront/input/
DIR elasticmapreduce/samples/cloudfront/logprocessor.jar
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-14.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-15.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-16.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-17.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-18.WxYz1234
```

```
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-19.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-20.WxYz1234
DIR elasticmapreduce/samples/cloudfront/code/cloudfront-loganalyzer.tgz
```

subcomando emrfs delete

[Opciones] para emrfs delete

Opción	Descripción	Obligatorio
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *METADATA_NAME* argumento, el valor predeterminado es `EmrFSMetadata`	No
s3://s3Path	La ruta al bucket de Amazon S3 que está siguiendo para una vista coherente. Los buckets se sincronizan de forma recursiva.	Sí
`-t` *TIME* | `--time` *TIME*	La fecha de vencimiento (interpretada con el argumento de unidad de tiempo). Todas las entradas de metadatos anteriores al *TIME* argumento se eliminan del depósito especificado.	
`-u` *UNIT* | `--time-unit` *UNIT*	La medida utilizada para interpretar el argumento de tiempo (nanosegundos, microsegundos, milisegundos, segundos, minutos, horas o días). Si no se especifica ningún argumento, el valor predeterminado es `days`.	
`--read-consumption` *READ_CONSUMPTION*	La cantidad solicitada de rendimiento de lectura disponible utilizada para la operación delete. Si no se especifica el *READ_CONSUMPTION* argumento, el valor predeterminado es`400`.	No

Opción	Descripción	Obligatorio
`--write-consumption` *WRITE_CONSUMPTION*	La cantidad solicitada de rendimiento de escritura disponible utilizada para la operación delete. Si no se especifica el *WRITE_CONSUMPTION* argumento, el valor predeterminado es `100`.	No

Example ejemplo de emrfs delete

El siguiente ejemplo elimina todos los objetos en un bucket de Amazon S3 desde los metadatos de seguimiento para una vista consistente.

```
$ emrfs delete s3://elasticmapreduce/samples/cloudfront
entries deleted: 11
```

subcomando emrfs import

[Opciones] para emrfs import

Opción	Descripción	Obligatorio
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *METADATA_NAME* argumento, el valor predeterminado es. `EmrFSMetadata`	No
s3://s3Path	La ruta al bucket de Amazon S3 que está siguiendo para una vista coherente. Los buckets se sincronizan de forma recursiva.	Sí
`--read-consumption` *READ_CONSUMPTION*	La cantidad solicitada de rendimiento de lectura disponible utilizada para la operación delete. Si no	No

Opción	Descripción	Obligatorio
	se especifica el *READ_CONSUMPTION* argumento, el valor predeterminado es `400`.	
`--write-consumption` *WRITE_CONSUMPTION*	La cantidad solicitada de rendimiento de escritura disponible utilizada para la operación delete. Si no se especifica el *WRITE_CONSUMPTION* argumento, el valor predeterminado es `100`.	No

Example ejemplo de emrfs import

El siguiente ejemplo importa todos los objetos en un bucket de Amazon S3 con los metadatos de seguimiento para una vista consistente. Todas las claves desconocidas se omiten.

```
$ emrfs import s3://elasticmapreduce/samples/cloudfront
```

subcomando emrfs sync

[Opciones] para emrfs sync

Opción	Descripción	Obligatorio
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* es el nombre de la tabla de metadatos de DynamoDB. Si no se proporciona el *METADATA_NAME* argumento, el valor predeterminado es. `EmrFSMetadata`	No
s3://s3Path	La ruta al bucket de Amazon S3 que está siguiendo para una vista coherente. Los buckets se sincronizan de forma recursiva.	Sí
		No

Opción	Descripción	Obligatorio
`--read-consumption` *READ_CONSUMPTION*	La cantidad solicitada de rendimiento de lectura disponible utilizada para la operación delete. Si no se especifica el *READ_CONSUMPTION* argumento, el valor predeterminado es `400`.	
`--write-consumption` *WRITE_CONSUMPTION*	La cantidad solicitada de rendimiento de escritura disponible utilizada para la operación delete. Si no se especifica el *WRITE_CONSUMPTION* argumento, el valor predeterminado es `100`.	No

Example ejemplo de comando emrfs sync

El siguiente ejemplo importa todos los objetos en un bucket de Amazon S3 con los metadatos de seguimiento para una vista consistente. Todas las claves desconocidas se eliminan.

```
$ emrfs sync s3://elasticmapreduce/samples/cloudfront
Synching samples/cloudfront                                  0 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/code/                            1 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/                                 2 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/input/                           9 added | 0 updated |
 0 removed | 0 unchanged
Done synching s3://elasticmapreduce/samples/cloudfront       9 added | 0 updated |
 1 removed | 0 unchanged
creating 3 folder key(s)
folders written: 3
```

subcomando emrfs read-sqs

[Opciones] para emrfs read-sqs

Opción	Descripción	Obligatorio

Opción	Descripción	Obligatorio
`-q` *QUEUE_NAME* | `--queue-name` *QUEUE_NAME*	*QUEUE_NAME* es el nombre de la cola de Amazon SQS configurada en. `emrfs-site.xml` El valor predeterminado es **`EMRFS-Inconsistency-<jobFlowId>`** .	Sí
`-o` *OUTPUT_FILE* | `--output-file` *OUTPUT_FILE*	*OUTPUT_FILE* es la ruta al archivo de salida en el sistema de archivos local del nodo maestro. Los mensajes leídos desde la cola se escriben en este archivo.	Sí

subcomando emrfs delete-sqs

[Opciones] para emrfs delete-sqs

Opción	Descripción	Obligatorio
`-q` *QUEUE_NAME* | `--queue-name` *QUEUE_NAME*	*QUEUE_NAME* es el nombre de la cola de Amazon SQS configurada en. `emrfs-site.xml` El valor predeterminado es **`EMRFS-Inconsistency-<jobFlowId>`** .	Sí

Envío de comandos de CLI EMRFS CLI como pasos

El siguiente ejemplo muestra cómo utilizar la `emrfs` utilidad en el nodo principal aprovechando la API AWS CLI o y la `command-runner.jar` para ejecutar el `emrfs` comando paso a paso. El ejemplo utiliza AWS SDK para Python (Boto3) para agregar un paso a un clúster que añade objetos en un bucket de Amazon S3 a la tabla de metadatos EMRFS predeterminada.

```
import boto3
from botocore.exceptions import ClientError


def add_emrfs_step(command, bucket_url, cluster_id, emr_client):
```

```
    """
    Add an EMRFS command as a job flow step to an existing cluster.

    :param command: The EMRFS command to run.
    :param bucket_url: The URL of a bucket that contains tracking metadata.
    :param cluster_id: The ID of the cluster to update.
    :param emr_client: The Boto3 Amazon EMR client object.
    :return: The ID of the added job flow step. Status can be tracked by calling
             the emr_client.describe_step() function.
    """
    job_flow_step = {
        "Name": "Example EMRFS Command Step",
        "ActionOnFailure": "CONTINUE",
        "HadoopJarStep": {
            "Jar": "command-runner.jar",
            "Args": ["/usr/bin/emrfs", command, bucket_url],
        },
    }

    try:
        response = emr_client.add_job_flow_steps(
            JobFlowId=cluster_id, Steps=[job_flow_step]
        )
        step_id = response["StepIds"][0]
        print(f"Added step {step_id} to cluster {cluster_id}.")
    except ClientError:
        print(f"Couldn't add a step to cluster {cluster_id}.")
        raise
    else:
        return step_id


def usage_demo():
    emr_client = boto3.client("emr")
    # Assumes the first waiting cluster has EMRFS enabled and has created metadata
    # with the default name of 'EmrFSMetadata'.
    cluster = emr_client.list_clusters(ClusterStates=["WAITING"])["Clusters"][0]
    add_emrfs_step(
        "sync", "s3://elasticmapreduce/samples/cloudfront", cluster["Id"], emr_client
    )


if __name__ == "__main__":
```

```
    usage_demo()
```

Puede utilizar el valor devuelto `step_id` para comprobar los registros para el resultado de la operación.

Autorización de acceso a los datos de EMRFS en Amazon S3

De forma predeterminada, el rol de EMR para EC2 determina los permisos para obtener acceso a los datos de EMRFS en Amazon S3. Las políticas de IAM asociadas a este rol se aplican independientemente del usuario o del grupo que realice la solicitud a través de EMRFS. El valor predeterminado es `EMR_EC2_DefaultRole`. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia).

A partir de la versión 5.10.0 de Amazon EMR, puede utilizar una configuración de seguridad para especificar los roles de IAM para EMRFS. Esto le permite personalizar los permisos para las solicitudes de EMRFS a Amazon S3 para los clústeres que tienen varios usuarios. Puede especificar roles de IAM diferentes para diferentes usuarios y grupos, y para ubicaciones de buckets de Amazon S3 diferentes basadas en el prefijo que tienen en Amazon S3. Cuando EMRFS realiza una solicitud a Amazon S3 que coincide con los usuarios, grupos o ubicaciones especificados, el clúster utiliza el rol correspondiente especificado en lugar del rol de EMR. EC2 Para obtener más información, consulte Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3.

Asimismo, si los requisitos de su solución de Amazon EMR no se satisfacen con los roles de IAM que EMRFS proporciona, puede definir una clase de proveedor de credenciales personalizada, que le permita personalizar el acceso a los datos de EMRFS en Amazon S3.

Creación de un proveedor de credenciales personalizado para los datos de EMRFS en Amazon S3

Para crear un proveedor de credenciales personalizado, debe implementar las clases AWSCredentialsProvider y Hadoop Configurable.

Para obtener una explicación detallada de este método, consulte Análisis seguro de datos desde otra AWS cuenta de con EMRFS en el blog de AWS macrodatos de. La entrada de blog incluye un tutorial que muestra el procesoend-to-end, desde la creación de los roles de IAM hasta el lanzamiento del clúster. También proporciona un ejemplo de código Java que implementa la clase del proveedor de credenciales personalizado.

Los pasos básicos son los siguientes:

Para especificar un proveedor de credenciales personalizado

1. Cree una clase de proveedor de credenciales personalizado compilada como un archivo JAR.
2. Ejecute un script como una acción de arranque para copiar el archivo JAR del proveedor de credenciales personalizado en la ubicación `/usr/share/aws/emr/emrfs/auxlib` del nodo principal del clúster. Para obtener más información acerca de las acciones de arranque, consulte (Opcional) Crear acciones de arranque para instalar software adicional.
3. Personalice la clasificación `emrfs-site` para especificar la clase que va a implementar en el archivo JAR. Para obtener más información acerca de cómo especificar objetos de configuración para personalizar aplicaciones, consulte Configuración de aplicaciones en la Guía de lanzamiento de Amazon EMR.

   En el siguiente ejemplo se muestra un comando `create-cluster` que lanza un clúster de Hive con parámetros de configuración comunes, y que también incluye:

   - Una acción de arranque que se ejecuta el script, *`copy_jar_file.sh`*, que se guarda en *`amzn-s3-demo-bucket`* en Amazon S3.
   - Una clasificación `emrfs-site` que especifica un proveedor de credenciales personalizado definido en el archivo JAR como *`MyCustomCredentialsProvider`*.

   > **Note**
   >
   > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --applications Name=Hive \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket/
copy_jar_file.sh","Name":"Custom action"}]' \
--ec2-attributes '{"KeyName":"MyKeyPair","InstanceProfile":"EMR_EC2_DefaultRole",\
"SubnetId":"subnet-xxxxxxxx","EmrManagedSlaveSecurityGroup":"sg-xxxxxxxx",\
"EmrManagedMasterSecurityGroup":"sg-xxxxxxxx"}' \
--service-role EMR_DefaultRole_V2 --enable-debugging --release-label emr-7.9.0 \
--log-uri 's3n://amzn-s3-demo-bucket/' --name 'test-awscredentialsprovider-emrfs' \
--instance-type=m5.xlarge --instance-count 3  \
--configurations '[{"Classification":"emrfs-site",\
```

```
"Properties":
{"fs.s3.customAWSCredentialsProvider":"MyAWSCredentialsProviderWithUri"},\
"Configurations":[]}]'
```

Administrar el AWS Security Token Service punto de conexión de predeterminado

EMRFS utiliza AWS Security Token Service (STS) para recuperar las credenciales de seguridad temporales con el fin de acceder a sus AWS recursos de. Las versiones anteriores de Amazon EMR envían todas las AWS STS solicitudes de a un único punto de conexión global en. `https://sts.amazonaws.com` Las versiones 5.31.0 y 6.1.0 y posteriores de Amazon EMR realizan solicitudes a puntos de conexión regionales de en su lugar. AWS STS Esto reduce la latencia y mejora la validez del token de sesión. Para obtener más información sobre los AWS STS puntos de enlace, consulte Administración AWS STS en una AWS región en la Guía del usuario.AWS Identity and Access Management

Cuando utiliza las versiones 5.31.0 y 6.1.0 y posteriores de Amazon EMR, puede anular el punto de conexión de AWS STS predeterminado. Para ello, debe cambiar la propiedad `fs.s3.sts.endpoint` en su configuración `emrfs-site`.

El siguiente AWS CLI ejemplo de establece el AWS STS punto de conexión de predeterminado utilizado por EMRFS como punto de conexión global.

```
aws emr create-cluster --release-label <emr-5.33.0> --instance-type m5.xlarge \
--emrfs Args=[fs.s3.sts.endpoint=https://sts.amazonaws.com]
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Como alternativa, puede crear un archivo de configuración JSON con el siguiente ejemplo y especificarlo con el argumento `--configurations` de `emr create-cluster`. Para obtener más información sobre `--configurations,`, consulte Referencia de comandos de AWS CLI .

```
[
  {
    "classification": "emrfs-site",
    "properties": {
      "fs.s3.sts.endpoint": "https://sts.amazonaws.com"
    }
  }
]
```

Especificación del cifrado de Amazon S3 mediante propiedades de EMRFS

> **Important**
>
> A partir de la versión 4.8.0 de Amazon EMR, puede utilizar las configuraciones de seguridad para aplicar la configuración de cifrado con mayor facilidad y con más opciones. Recomendamos que utilice las configuraciones de seguridad. Para obtener más información, consulte Configuración del cifrado de datos. Las instrucciones de la consola que se describen en esta sección están disponibles para las versiones anteriores a la 4.8.0. Si utiliza la AWS CLI para configurar el cifrado Amazon S3 tanto en la configuración del clúster como en una configuración de seguridad en las versiones posteriores, la configuración de seguridad anula la la configuración del clúster.

Al crear un clúster, puede especificar el cifrado del servidor (SSE) o el cifrado del cliente (CSE) para los datos de EMRFS en Amazon S3 usando la consola o mediante las propiedades de clasificación a través de la o `emrfs-site` del SDK de EMR. AWS CLI SSE y de CSE de Amazon S3 son mutuamente excluyentes, puede elegir uno de ellos, pero no ambos.

Para obtener AWS CLI instrucciones, consulte la sección correspondiente a su tipo de cifrado a continuación.

Para especificar las opciones de cifrado de EMRFS mediante la AWS Management Console

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.

3. Para Release (Versión), elija 4.7.2 o anterior.
4. Elija otras opciones para Software and Steps (Software y pasos) adecuadas para su aplicación y, a continuación, elija Next (Siguiente).
5. En los paneles Hardware y General Cluster Settings (Configuración general del clúster), elija opciones adecuadas para su aplicación.
6. En el panel Security (Seguridad), en Authentication and encryption (Autenticación y cifrado), seleccione la opción S3 Encryption (with EMRFS) (Cifrado de S3 (con EMRFS)).

   > **Note**
   >
   > La opción Cifrado del servidor de S3 con KMS Key Management (SSE-KMS) no está disponible cuando se utiliza Amazon EMR versión 4.4 o anterior.

   - Si elige una opción que utiliza AWS Key Management, elija un identificador en el campo ID de clave de AWS KMS. Para obtener más información, consulte Utilización de AWS KMS keys para el cifrado de EMRFS.
   - Si elige S3 client-side encryption with custom materials provider (Cifrado del lado del cliente de S3 con el proveedor de materiales de cifrado personalizado), rellene los campos Class name (Nombre de la clase) y JAR location (Ubicación de JAR). Para obtener más información, consulte Cifrado del cliente de Amazon S3.
7. Elija las demás opciones que sean necesarias para su aplicación y, a continuación, elija Create Cluster (Crear clúster).

Utilización de AWS KMS keys para el cifrado de EMRFS

La clave de AWS KMS cifrado de debe estar creada en la misma región que su instancia de clúster de Amazon EMR y los buckets de Amazon S3 utilizados con EMRFS. Si la clave que especifica está en una cuenta diferente de la que usa para configurar un clúster, debe especificar la clave mediante su ARN.

El rol del perfil de EC2 instancia de Amazon debe tener permisos para usar la clave de KMS que especifique. El rol predeterminado del perfil de instancia en Amazon EMR es `EMR_EC2_DefaultRole`. Si usa un rol diferente para el perfil de instancia o usa roles de IAM para las solicitudes de EMRFS a Amazon S3, asegúrese de agregar cada rol como usuario clave, según corresponda. Esto proporciona al rol los permisos para utilizar la clave de KMS. Para obtener

más información, consulte Uso de políticas de claves en la Guía para desarrolladores de AWS Key Management Service y Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3.

Puede utilizar la AWS Management Console para agregar su perfil de instancia o el perfil de EC2 instancia a la lista de usuarios clave de la clave de KMS especificada, o bien puede usar la AWS CLI o un AWS SDK para adjuntar una política de claves adecuada.

Tenga en cuenta que Amazon EMR solo admite claves de KMS simétricas. No se puede utilizar una clave KMS asimétrica para cifrar datos en reposo en un clúster de Amazon EMR. Para obtener ayuda para determinar si una clave de KMS es simétrica o asimétrica, consulte Identificación de claves de KMS simétricas y asimétricas.

El siguiente procedimiento describe cómo agregar el perfil de instancia de Amazon EMR predeterminado, `EMR_EC2_DefaultRole` como un usuario clave con la AWS Management Console. Se supone que ya ha creado una clave de KMS. Para crear una nueva clave de KMS, consulte Creación de claves en la Guía para desarrolladores de AWS Key Management Service .

Para agregar el perfil de EC2 instancia de Amazon EMR a la lista de usuarios de claves de cifrado

1. Inicie sesión en la consola AWS Key Management Service (AWS KMS) AWS Management Console y ábrala en https://console.aws.amazon.com/kms.
2. Para cambiar la Región de AWS, utilice el Selector de regiones ubicado en la esquina superior derecha de la página.
3. Seleccione el alias de la clave de KMS que desee modificar.
4. En la página de detalles de la clave, en Key Users (Usuarios de claves), seleccione Add (Añadir).
5. En el cuadro de diálogo Add key users (Añadir usuarios clave), seleccione el rol adecuado. El nombre del rol predeterminado es `EMR_EC2_DefaultRole`.
6. Elija Agregar.

Cifrado del servidor de Amazon S3

Cuando configura el cifrado del servidor de Amazon S3, Amazon S3 cifra los datos del objeto a medida que escribe los datos en el disco y descifra los datos cuando se accede. Para obtener más información sobre SSE, consulte Protección de los datos con el cifrado del servidor en la Guía del usuario de Amazon Simple Storage Service.

Puede elegir entre dos sistemas de administración de claves distintos al especificar SSE en Amazon EMR:

- SSE-S3: Amazon S3 administra las claves en su nombre.
- SSE-KMS: utiliza una AWS KMS key para configurar políticas adecuadas para Amazon EMR. Para obtener información adicional sobre los requisitos de las claves de Amazon EMR, consulte Uso de AWS KMS keys para el cifrado.

SSE con claves proporcionadas por el cliente (SSE-C) no está disponible para su uso con Amazon EMR.

Para crear un clúster con el SSE-S3 habilitado mediante el AWS CLI

- Escriba el siguiente comando:

```
aws emr create-cluster --release-label emr-4.7.2 or earlier \
--instance-count 3 --instance-type m5.xlarge --emrfs Encryption=ServerSide
```

También puede habilitar SSE-S3 configurando fs.s3. enableServerSideLa propiedad de cifrado es verdadera en las propiedades. `emrfs-site` Consulte el ejemplo de SSE-KMS a continuación y omita la propiedad para el ID de clave.

Para crear un clúster con el SSE-KMS habilitado mediante el AWS CLI

> **Note**
>
> SSE-KMS solo está disponible en la versión 4.5.0 y posteriores de Amazon EMR.

- Escriba el siguiente AWS CLI comando para crear un clúster con SSE-KMS, donde *keyID* hay, por ejemplo AWS KMS key, un: *a4567b8-9900-12ab-1234-123a45678901*

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3 \
--instance-type m5.xlarge --use-default-roles \
--emrfs Encryption=ServerSide,Args=[fs.s3.serverSideEncryption.kms.keyId=keyId]
```

--O BIEN--

Escriba el siguiente AWS CLI comando utilizando la `emrfs-site` clasificación y proporcione un archivo JSON de configuración con el contenido que se muestra de forma similar al `myConfig.json` del ejemplo siguiente:

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3
 --instance-type m5.xlarge --applications Name=Hadoop --configurations file://
myConfig.json --use-default-roles
```

Ejemplo de contenido de myConfig.json:

```
[
  {
    "Classification":"emrfs-site",
    "Properties": {
       "fs.s3.enableServerSideEncryption": "true",
       "fs.s3.serverSideEncryption.kms.keyId":"a4567b8-9900-12ab-1234-123a45678901"
    }
  }
]
```

Propiedades de configuración para SSE-S3 y SSE-KMS

Estas propiedades se pueden configurar mediante la clasificación de configuración `emrfs-site`. SSE-KMS solo está disponible en la versión 4.5.0 y posteriores de Amazon EMR.

Propiedad	Valor predeterminado	Descripción
`fs.s3.enableServerSideEncryption`	**`false`**	Cuando se establecen en **`true`**, los objetos almacenados en Amazon S3 se cifran mediante el cifrado del servidor. Si no se especifica ninguna clave, se utiliza SSE-S3.

Propiedad	Valor predeterminado	Descripción
`fs.s3.serverSideEncryption.kms.keyId`	**n/a**	Especifica un identificador AWS KMS clave o ARN. Si se especifica una clave, se utiliza SSE-KMS.

Cifrado del cliente de Amazon S3

Con el cifrado del cliente de Amazon S3, el proceso de cifrado y descifrado de Amazon S3 se produce en el cliente de EMRFS en su clúster. Los objetos se cifran antes de cargarlos en Amazon S3 y se descifran después de que se descarguen. El proveedor que especifique proporciona la clave de cifrado que utiliza el cliente. El cliente puede usar claves proporcionadas por AWS KMS (CSE-KMS) o una clase de Java personalizada que proporciona la clave raíz del cliente (CSE-C). Los detalles de cifrado son ligeramente diferentes entre CSE-KMS y CSE-C, en función del proveedor especificado y de los metadatos del objeto que se descifra o se cifra. Para obtener más información sobre estas diferencias, consulte Protección de los datos con el cifrado del cliente en la Guía del usuario de Amazon Simple Storage Service.

> **Note**
>
> El CSE de Amazon S3 solo garantiza que los datos de EMRFS intercambiados con Amazon S3 se cifren; no se cifran todos los datos en volúmenes de instancias de clúster. Además, ya que Hue no utiliza EMRFS, los objetos que Hue S3 File Browser escribe en Amazon S3 no se cifran.

Para especificar CSE-KMS para datos de EMRFS en Amazon S3 mediante la AWS CLI

- Escriba el siguiente comando y *MyKMSKeyID* sustitúyalo por el ID de clave o el ARN de la clave de KMS que se debe utilizar:

```
aws emr create-cluster --release-label emr-4.7.2 or earlier
--emrfs Encryption=ClientSide,ProviderType=KMS,KMSKeyId=MyKMSKeyId
```

Creación de un proveedor de claves personalizadas

Según el tipo de cifrado que utilice al crear un proveedor de claves personalizadas, la aplicación también debe implementar diferentes EncryptionMaterialsProvider interfaces. Ambas interfaces están disponibles en la versión 1.11.0 y posteriores del AWS SDK for Java.

- Para implementar el cifrado de Amazon S3, utilice com.amazonaws.services.s3.model. EncryptionMaterialsProvider interfaz.
- Para implementar el cifrado del disco local, utilice com.amazonaws.services.elasticmapreduce.spi.security. EncryptionMaterialsProvider interfaz.

Puede utilizar cualquier estrategia para proporcionar materiales de cifrado para la implementación. Por ejemplo, podría elegir proporcionar materiales de cifrado estáticos o integrar con un sistema de administración de claves más complejo.

Si utiliza el cifrado de Amazon S3, debe utilizar los algoritmos de cifrado AES/GCM/NoPaddingpara los materiales de cifrado personalizados.

Si utiliza el cifrado de disco local, el algoritmo de cifrado que se utilizará para los materiales de cifrado personalizados varía según la versión del EMR. Para Amazon EMR 7.0.0 y versiones anteriores, debe usar. AES/GCM/NoPadding Para Amazon EMR 7.1.0 y versiones posteriores, debe usar AES.

La EncryptionMaterialsProvider clase obtiene los materiales de cifrado por contexto de cifrado. Amazon EMR rellena el contexto de cifrado en tiempo de ejecución para ayudar al intermediario a determinar qué materiales de cifrado debe devolver.

Example Ejemplo: uso de un proveedor de claves de cifrado personalizadas para el cifrado de Amazon S3 con EMRFS

Cuando Amazon EMR recupera los materiales de cifrado desde la EncryptionMaterialsProvider clase para realizar el cifrado, EMRFS rellena opcionalmente el argumento materialsDescription con dos campos: el URI de Amazon S3 para el objeto y el del clúster, que puede utilizar JobFlowId la clase para devolver materiales de cifrado de manera selectiva. EncryptionMaterialsProvider

Por ejemplo, el proveedor podría devolver claves distintas para diferentes prefijos URI de Amazon S3. Se trata de la descripción de los materiales de cifrado devuelta que se almacena finalmente con el objeto de Amazon S3 en lugar del valor materialsDescription que genera EMRFS y se transfiere

al proveedor. Al descifrar un objeto de Amazon S3, la descripción de materiales de cifrado se transfiere a la EncryptionMaterialsProvider clase, de modo que puede, nuevamente, devolver de forma selectiva la clave coincidente para descifrar el objeto.

A continuación, se proporciona una implementación de EncryptionMaterialsProvider referencia. Otro proveedor personalizado, EMRFSRSAEncryptionMaterialsProvider, está disponible en GitHub.

```
import com.amazonaws.services.s3.model.EncryptionMaterials;
import com.amazonaws.services.s3.model.EncryptionMaterialsProvider;
import com.amazonaws.services.s3.model.KMSEncryptionMaterials;
import org.apache.hadoop.conf.Configurable;
import org.apache.hadoop.conf.Configuration;

import java.util.Map;

/**
 * Provides KMSEncryptionMaterials according to Configuration
 */
public class MyEncryptionMaterialsProviders implements EncryptionMaterialsProvider,
 Configurable{
  private Configuration conf;
  private String kmsKeyId;
  private EncryptionMaterials encryptionMaterials;

  private void init() {
    this.kmsKeyId = conf.get("my.kms.key.id");
    this.encryptionMaterials = new KMSEncryptionMaterials(kmsKeyId);
  }

  @Override
  public void setConf(Configuration conf) {
    this.conf = conf;
    init();
  }

  @Override
  public Configuration getConf() {
    return this.conf;
  }

  @Override
  public void refresh() {
```

```
  }

  @Override
  public EncryptionMaterials getEncryptionMaterials(Map<String, String>
 materialsDescription) {
    return this.encryptionMaterials;
  }

  @Override
  public EncryptionMaterials getEncryptionMaterials() {
    return this.encryptionMaterials;
  }
}
```

Especificación de un proveedor de materiales personalizado mediante la AWS CLI

Para usar la AWS CLI, transfiera los argumentos `Encryption`, `ProviderType`, `CustomProviderClass` y `CustomProviderLocation` a la opción `emrfs`.

```
aws emr create-cluster --instance-type m5.xlarge --release-label emr-4.7.2 or earlier
 --emrfs Encryption=ClientSide,ProviderType=Custom,CustomProviderLocation=s3://amzn-s3-
demo-bucket/myfolder/provider.jar,CustomProviderClass=classname
```

Al definir `Encryption` en `ClientSide`, se habilita el cifrado del cliente, `CustomProviderClass` es el nombre del objeto `EncryptionMaterialsProvider` y `CustomProviderLocation` es la ubicación local o de Amazon S3 desde la que Amazon EMR copia `CustomProviderClass` en cada nodo del clúster y lo coloca en el classpath.

Especificación de un proveedor de materiales personalizado mediante un SDK

Para utilizar un SDK, puede establecer la propiedad `fs.s3.cse.encryptionMaterialsProvider.uri` para descargar en cada nodo del clúster la clase `EncryptionMaterialsProvider` personalizada que se almacena en Amazon S3. Esto se configura en el archivo `emrfs-site.xml` junto con CSE habilitado y la ubicación correcta del proveedor personalizado.

Por ejemplo, al AWS SDK para Java usar RunJobFlowRequest, el código podría verse así:

```
<snip>
  Map<String,String> emrfsProperties = new HashMap<String,String>();
```

```
      emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider.uri","s3://amzn-s3-
demo-bucket/MyCustomEncryptionMaterialsProvider.jar");
      emrfsProperties.put("fs.s3.cse.enabled","true");
      emrfsProperties.put("fs.s3.consistent","true");

 emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider","full.class.name.of.EncryptionMate

  Configuration myEmrfsConfig = new Configuration()
      .withClassification("emrfs-site")
      .withProperties(emrfsProperties);

  RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("Custom EncryptionMaterialsProvider")
   .withReleaseLabel("emr-7.9.0")
   .withApplications(myApp)
   .withConfigurations(myEmrfsConfig)
   .withServiceRole("EMR_DefaultRole_V2")
   .withJobFlowRole("EMR_EC2_DefaultRole")
   .withLogUri("s3://myLogUri/")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myEc2Key")
    .withInstanceCount(2)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m5.xlarge")
    .withSlaveInstanceType("m5.xlarge")
   );

  RunJobFlowResult result = emr.runJobFlow(request);
</snip>
```

Personalización de EncryptionMaterialsProvider con argumentos

Es posible que necesite transferir argumentos directamente al proveedor. Para ello, puede utilizar la clasificación de configuración `emrfs-site` con argumentos personalizadas definidos como propiedades. A continuación se muestra un ejemplo de configuración que se guarda como un archivo, `myConfig.json`:

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "myProvider.arg1":"value1",
     "myProvider.arg2":"value2"
```

```
      }
    }
 ]
```

Con el `create-cluster` comando de AWS CLI, puede usar la `--configurations` opción para especificar el archivo, tal como se muestra a continuación:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge
 --instance-count 2 --configurations file://myConfig.json --emrfs
 Encryption=ClientSide,CustomProviderLocation=s3://amzn-s3-demo-bucket/myfolder/
myprovider.jar,CustomProviderClass=classname
```

Configuración del soporte para la versión 2 de EMRFS S3EC

Las versiones 1.11.837 y posteriores del SDK para Java de S3, se ha introducido la versión de soporte 2 del cliente de cifrado (S3EC V2) con varias mejoras de seguridad. Para obtener más información, consulte la publicación en el blog de S3: Updates to the Amazon S3 encryption client. Consulte también la Migración del cifrado del cliente de Amazon S3 en la Guía para AWS SDK para Java desarrolladores de.

La versión 1 del cliente de cifrado todavía está disponible en el SDK por motivos de compatibilidad con versiones anteriores. De forma predeterminada, EMRFS utilizará la versión 1 de S3EC para cifrar y descifrar los objetos de S3 si CSE está habilitado.

EMRFS no puede descifrar los objetos de S3 cifrados con la versión 2 de S3EC en un clúster de EMR cuya versión de lanzamiento sea anterior a emr-5.31.0 (emr-5.30.1 y anteriores, emr-6.1.0 y anteriores).

Example Configuración de EMRFS para usar la versión 2 de S3EC

A fin de configurar EMRFS para que utilice la versión 2 de S3EC, agregue la siguiente configuración:

```
{
  "Classification": "emrfs-site",
  "Properties": {
    "fs.s3.cse.encryptionV2.enabled": "true"
  }
}
```

emrfs-site.xml Propiedades para el cifrado del cliente de Amazon S3

Propiedad	Valor predeterminado	Descripción
`fs.s3.cse.enabled`	**`false`**	Cuando se establece en **`true`**, los objetos de EMRFS almacenados en Amazon S3 se cifran mediante el cifrado del cliente.
`fs.s3.cse.encryptionV2.enabled`	**`false`**	Cuando se establece en `true`, EMRFS utiliza la versión 2 del cliente de cifrado de S3 para cifrar y descifrar objetos en S3. Está disponible en la versión 5.31.0 y posteriores de EMR.
`fs.s3.cse.encryptionMaterialsProvider.uri`	**N/A**	Se aplica cuando se utilizan materiales de cifrado personalizados. URI de Amazon S3 en el que se encuentra el JAR con `EncryptionMaterialsProvider` . Al proporcionar este URI, Amazon EMR descarga automáticamente el JAR a todos los nodos del clúster.
`fs.s3.cse.encryptionMaterialsProvider`	**N/A**	La ruta de clases `EncryptionMaterialsProvider` utilizada con el cifrado del lado del cliente. Cuando utilice CSE-KMS, especifique `com.amazo`

Propiedad	Valor predeterminado	Descripción
		`n.ws.emr.hadoop.fs.cse.KMSEncryptionMaterialsProvider` .
`fs.s3.cse.materialsDescription.enabled`	**`false`**	Cuando se establece en`true`, rellena el argumento materials Description de los objetos cifrados con el URI de Amazon S3 del objeto y con. JobFlowId Establézcalo en `true` cuando se utilicen materiales de cifrado personalizados.
`fs.s3.cse.kms.keyId`	**`N/A`**	Se aplica cuando se utiliza CSE-KMS. Es el valor del ARN KeyId o alias de la clave de KMS; utilizada para el cifrado.

Propiedad	Valor predeterminado	Descripción
`fs.s3.cse.cryptoStorageMode`	**`ObjectMetadata`**	Modo de almacenamiento de Amazon S3. De forma predeterminada, la descripción de la información de cifrado se almacena en los metadatos de objetos. También puede almacenar la descripción en un archivo de instrucciones. Los valores válidos son ObjectMetadata y. InstructionFile Para obtener más información adicional, consulte Cifrado de datos del cliente con AWS SDK para Java y Amazon S3.

Cifrado del lado del cliente de Amazon S3 con S3A

A partir de la versión 7.6.0 de Amazon Elastic Map Reduce (EMR), el conector del sistema de archivos S3A ahora admite el cifrado del lado del cliente de Amazon S3. Esto significa que el cifrado y el descifrado de los datos de Amazon S3 se realizan directamente en el cliente S3A de su clúster informático. Al utilizar esta función, los archivos se cifran automáticamente antes de cargarlos en Amazon S3 y se descifran al descargarlos. Para obtener información detallada sobre la metodología de cifrado y su implementación, los usuarios pueden consultar la sección Protección de datos mediante el cifrado del lado del cliente en la Guía del usuario de Amazon Simple Storage Service.

Al habilitar el cifrado del lado del cliente (CSE) con S3A en Amazon EMR, tiene dos opciones de sistema de administración de claves:

- CSE-KMS: este enfoque utiliza una clave AWS Key Management Service (KMS) configurada con políticas diseñadas específicamente para Amazon EMR. Para obtener información detallada sobre los requisitos clave, consulte la documentación sobre el uso de AWS KMS claves para el cifrado.
- CSE-CUSTOM: este método le permite integrar una clase Java personalizada que proporciona la clave raíz del lado del cliente responsable de cifrar y descifrar los datos.

> **Note**
>
> El cifrado del lado del cliente S3A en EMR es intrínsecamente compatible con el cifrado del lado del cliente de EMRFS, lo que significa que los objetos cifrados con EMRFS CSE se pueden leer a través del S3A CSE.

Temas

- Configuración CSE-KMS
- Configuración: CSE-CUSTOM
- Propiedades del cifrado del lado del cliente de Amazon S3 con S3A

Configuración CSE-KMS

Puede habilitar el cifrado del lado del cliente mediante AWS KMS (CSE-KMS) en dos ámbitos principales:

- El primero es la configuración de todo el clúster:

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-KMS",
       "fs.s3a.encryption.key":"${KMS_KEY_ID}",
    }
  }
]
```

> **Note**
>
> Si la región AWS KMS clave es diferente de la región S3 Bucket/EMR, debe establecer la siguiente configuración adicional:. `fs.s3a.encryption.cse.kms.region=${KMS_REGION}`

- La segunda es la configuración específica del trabajo o la aplicación. El CSE-KMS se puede configurar para una aplicación Spark específica de la siguiente manera:

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-KMS --conf
 spark.hadoop.fs.s3a.encryption.key=${KMS_KEY_ID}
```

Configuración: CSE-CUSTOM

Para utilizar CSE-CUSTOM, debe crear un proveedor de claves personalizado mediante la implementación de la interfaz Keyring. A continuación, se muestra un ejemplo de implementación:

```
public class CustomKeyring implements Keyring {
  public CustomKeyring()  {
    // custom code
  }

  @Override
  public EncryptionMaterials onEncrypt(EncryptionMaterials encryptionMaterials) {
    // custom code
  }
```

```
  @Override
  public DecryptionMaterials onDecrypt(DecryptionMaterials decryptionMaterials,
      ListEncryptedDataKey list) {
    // custom code
  }
```

Puede habilitar las claves personalizadas de cifrado del lado del cliente (CSE-CUSTOM) en dos ámbitos principales:

- El primer ámbito es la configuración de todo el clúster:

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-CUSTOM",
       "fs.s3a.cse.customKeyringProvider.uri":"S3 path of custom jar",
       "fs.s3a.encryption.cse.custom.keyring.class.name":"fully qualified class name"
    }
  }
]
```

- El segundo es la configuración específica del trabajo o de la aplicación. El CSE-CUSTOM se puede configurar para una aplicación Spark específica de la siguiente manera:

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-CUSTOM --conf
 spark.hadoop.fs.s3a.encryption.cse.custom.keyring.class.name=fully qualified class
 name
```

> **Note**
>
> Asegúrese de que el contenedor personalizado necesario para generar las claves de cifrado y descifrado esté presente en la ruta de clases.

Propiedades del cifrado del lado del cliente de Amazon S3 con S3A

Para configurar el cifrado del lado del cliente con S3A, hay varias propiedades de configuración que se deben establecer en los ajustes de core-site.xml. Para obtener más información sobre los ajustes de configuración personalizados, consulte Configurar aplicaciones.

Propiedad	Valor predeterminado	Descripción
fs.s3a.encryption.algorithm	N/A	Cuando se establece en `CSE-KMS` o`CSE-CUSTOM` , los objetos almacenados en Amazon S3 se cifran mediante el cifrado del lado del cliente.
fs.s3a.encryption.key	N/A	Se aplica cuando se usa. `CSE-KMS` El valor del ARN KeyId o el alias de la clave KMS utilizada para el cifrado.
fs.s3a.encryption.cse.kms.region	N/A	Se aplica cuando se usa. `CSE-KMS` La región en la que se genera la AWS KMS clave. De forma predeterminada, la región KMS se establece en los mismos valores que la región del clúster EMR o del bucket S3.
fs.s3a.encryption.cse.custom.keyring.class.name	N/A	Se aplica cuando se usa. `CSE-KMS` El nombre de clase completo del proveedor de claves personalizadas.
fs.s3a.cse. customKeyringProvider.uri	N/A	Se aplica cuando se usa`CSE-CUSTOM` . El URI de Amazon S3 donde se encuentra el JAR con la implementación personalizada de Keyring. Al proporcionar este URI, Amazon EMR descarga automáticamente el JAR a todos los nodos del clúster.

Propiedad	Valor predeterminado	Descripción
fs.s3a.encryption.cse.v1.compatibility.enabled	`'true'`	Esto proporciona compatibilidad con versiones anteriores de clientes del SDK más antiguos, como el que se utilizaba con EMRFS. Desactívela cuando no exista tal dependencia para obtener un mejor rendimiento.

CloudWatch Agente de Amazon

El CloudWatch agente de Amazon en Amazon EMR es una herramienta que puede supervisar las EC2 instancias de Amazon en su clúster de EMR. Puede almacenar y ver las métricas que recopila con el CloudWatch agente CloudWatch. Para obtener más información sobre el CloudWatch agente, consulta la Guía del CloudWatch usuario de Amazon.

> **Note**
>
> Se le aplicarán cargos adicionales si utiliza otros Servicios de AWS para publicar, consultar o ver las métricas de los CloudWatch agentes de Amazon. Para obtener más información, consulte las páginas siguientes sobre precios.
>
> - CloudWatch Precios de Amazon
> - Amazon Managed Service para precios de Prometheus
> - Precios de Amazon Managed Grafana
>
> Estas métricas son independientes de las métricas gratuitas que Amazon EMR proporciona en el espacio de `AWS/ElasticMapReduce` nombres de. CloudWatch Para obtener más información acerca de las métricas que el CloudWatch agente no publica, consulte Monitorización de métricas de Amazon EMR con. CloudWatch

Con Amazon EMR 7.0 y versiones posteriores, puede instalar una versión personalizada del CloudWatch agente de Amazon en un clúster para recopilar métricas de su clúster de EMR. CloudWatchAgent es compatible con los clústeres de roles en tiempo de ejecución para EMR 7.6 y versiones posteriores. El agente recopila las métricas de todos los nodos en el clúster, las reúne en el nodo principal y las publica desde ese nodo en la nube.

Puede instalar el agente al crear un nuevo clúster en la consola o al utilizar la API `create-cluster`. Para obtener más información, consulte Creación de un clúster de EMR que utilice el agente de Amazon CloudWatch . De forma predeterminada, un clúster de Amazon EMR que ejecuta la serie 7.x publica 34 métricas a nivel de sistema CloudWatch en intervalos de 60 segundos, pero puede configurar el agente para que publique métricas diferentes. Otra opción es que publique métricas en Amazon Managed Service para Prometheus y puede elegir qué métricas desea que publique también. En diferentes casos de uso y configuraciones, puede configurar al agente para ver

y solicitar las métricas en la CloudWatch consola, en Amazon Managed Grafana o a través del para Prometheus CloudWatch o Amazon Managed APIs Service para Prometheus. Se aplican cargos de Servicios de AWS cuando se utilizan para almacenar y consultar métricas.

El CloudWatch agente de Amazon EMR puede publicar métricas del sistema y métricas de JMX para los siguientes servicios en su clúster de Amazon EMR.

- Hadoop DataNode
- Hadoop NameNode
- Hilo NodeManager
- Hilo ResourceManager
- HBase Master: solo Amazon EMR 7.1+
- HBase RegionServer — solo Amazon EMR 7.1+
- HBase ThriftServer — solo Amazon EMR 7.1+

Para obtener más información acerca de las métricas disponibles y de cómo configurar el CloudWatch agente en Amazon EMR, consulte Configuración del CloudWatch agente para Amazon EMR.

En la siguiente tabla, se muestra la versión AmazonCloudWatchAgent incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes con los que se instala Amazon EMR. AmazonCloudWatchAgent

Para ver la versión de los componentes instalados AmazonCloudWatchAgent en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

AmazonCloudWatchAgent información de versión 7.9.0

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
emr7.9.0	AmazonCloudWatchAgent 1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
		library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

Temas

- Creación de un clúster de EMR que utilice el agente de Amazon CloudWatch
- Métricas predeterminadas para el CloudWatch agente con Amazon EMR
- Configuración del CloudWatch agente para Amazon EMR
- Consideraciones y limitaciones
- CloudWatch historial de lanzamientos del agente

Creación de un clúster de EMR que utilice el agente de Amazon CloudWatch

Los procedimientos de esta sección describen los pasos para crear un clúster en Amazon EMR con el CloudWatch agente de Amazon desde la AWS Management Console y la. AWS CLI

Temas

- Permisos de IAM necesarios para el agente CloudWatch
- Punto de conexión CloudWatch del agente requerido
- Creación de un clúster de EMR

Permisos de IAM necesarios para el agente CloudWatch

El CloudWatch agente necesita el `cloudwatch:PutMetricData` permiso AWS Identity and Access Management (IAM) en el perfil de EC2 instancia de Amazon para Amazon EMR. El rol

predeterminado de Amazon EMR ya tiene este permiso. Puede crear el rol predeterminado desde aquí. AWS CLI `aws emr create-default-roles` Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia) en la Guía de administración de Amazon EMR.

El siguiente ejemplo de política de IAM incluye el permiso `cloudwatch:PutMetricData`:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "cloudwatch:PutMetricData",
      "Resource": "*"
    }
  ]
}
```

Punto de conexión CloudWatch del agente requerido

Para publicar métricas en un clúster de EMR en una subred privada, cree un punto de conexión de CloudWatch agente y asócielo a la VPC en la que se encuentra la subred privada. CloudWatch

Para obtener más información sobre los CloudWatch puntos de enlace de cada uno Región de AWS, consulta los CloudWatch puntos de enlace y las cuotas de Amazon en la Guía de referencia AWS general.

Creación de un clúster de EMR

Una vez que haya configurado los permisos y el punto de conexión obligatorios para utilizarlos con el CloudWatch agente, utilice la AWS Management Console o la AWS CLI para crear un nuevo clúster con el agente instalado.

Console

Para crear un clúster con el CloudWatch agente de Amazon desde la consola

1. Vaya a la consola de Amazon EMR.
2. Elija Create cluster.
3. En Nombre y aplicaciones, elija la versión de Amazon EMR 7.0.0 o posterior.

4. En Paquete de aplicaciones, seleccione el paquete o las aplicaciones que desea instalar en el clúster e incluya el CloudWatch agente en la selección.
5. Continúe con la creación del clúster que se ajuste a las necesidades del caso de uso.

AWS CLI

En AWS CLI, puedes añadir un CloudWatch agente de Amazon a un clúster con el `--applications` parámetro for`create-cluster`.

Para crear un clúster con Amazon CloudWatch Agent desde AWS CLI

- Al crear un clúster, utilice un comando similar al siguiente para incluir el CloudWatch agente de Amazon. *myKey*Sustituya por el nombre de su EC2 key pair.

```
aws emr create-cluster --name "Spark cluster with CloudWatch agent" \
--release-label emr-7.0.0 \
--applications Name=Spark Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--instance-count 3 --use-default-roles
```

Para obtener más información sobre cómo utilizar Amazon EMR con la AWS CLI, consulte la Referencia de AWS CLI comandos de la.

Métricas predeterminadas para el CloudWatch agente con Amazon EMR

Al instalar el CloudWatch agente de Amazon en Amazon EMR, la configuración predeterminada publica las siguientes métricas del sistema para todas las instancias del clúster, a menos que configure el agente de forma diferente. Para ver las definiciones de cada métrica, consulta las métricas recopiladas por el CloudWatch agente en la Guía del CloudWatch usuario de Amazon.

CPU

Métricas de CPU

- `cpu_usage_active`
- `cpu_usage_guest`

- cpu_usage_guest_nice
- cpu_usage_idle
- cpu_usage_iowait
- cpu_usage_irq
- cpu_usage_nice
- cpu_usage_softirq
- cpu_usage_steal
- cpu_usage_system
- cpu_usage_user

Disk

Métricas de disco

- disk_free
- disk_total
- disk_used
- disk_used_percent

Memory

Métricas de memoria

- mem_active
- mem_available
- mem_available_percent
- mem_free
- mem_inactive
- mem_total
- mem_used
- mem_used_percent
- mem_buffered
- mem_cached

Network IO

Métricas de E/S de red

- `net_bytes_recv`
- `net_bytes_sent`
- `net_packets_recv`
- `net_packets_sent`

Process

Métricas de proceso

- `processes_running`
- `processes_total`

Swap

Métricas de intercambio

- `swap_free`
- `swap_used`
- `swap_used_percent`

De forma predeterminada, el agente publica todas las métricas CloudWatch en el espacio de nombres `CWAgent` personalizado y en el esquema`cluster.id`,,`instance.id`,`node.type`. `service.name` Realice los siguientes pasos para ver estas métricas predeterminadas que el CloudWatch agente publica para Amazon EMR:

1. Vaya a la CloudWatch consola de Amazon.
2. Elija la pestaña Métricas y, a continuación, Todas las métricas.
3. En Espacios de nombres personalizados, seleccione **`CWAgent`**. A continuación, seleccione el esquema: **`cluster.id`**, **`instance.id`**, **`node.type`**, **`service.name`**.
4. Continúe con la consulta de las métricas desde la CloudWatch interfaz según lo requiera el caso de uso.

Amazon EMR asocia las siguientes etiquetas a cada métrica: `["cluster.id", "instance.id', "node.type", "service.name"]`. Tenga en cuenta lo siguiente con respecto a estas etiquetas:

- El valor de la etiqueta `cluster.id` es el ID del clúster de EMR que produjo la métrica. Un ejemplo de valor de la etiqueta `cluster.id` es: `j-123456789ABC`.
- El valor de la etiqueta `instance.id` es el ID de la instancia del clúster de EMR que produjo la métrica. Un ejemplo de valor de la etiqueta `instance.id` es: `i-01bcf5f140f3355777`.
- El valor de la etiqueta `node.type` representa el tipo de nodo de la instancia del clúster que produjo la métrica. El valor puede ser `master`, `core` o `task`.
- El valor de la etiqueta `service.name` es el nombre del servicio que produjo la métrica. El valor predeterminado del `service.name` de las métricas predeterminadas del sistema es `system`.

> **Note**
>
> La versión 7.0.0 de Amazon EMR utiliza un esquema de etiquetas de métricas ligeramente diferente. `cluster.id` es `jobflow.id` y `node.type` no está disponible. `instance.id` y `service.name` están disponibles.

Configuración del CloudWatch agente para Amazon EMR

La versión 7.0.0 y posteriores de Amazon EMR incluyen el agente de Amazon. CloudWatch En las siguientes secciones, se describe cómo configurar el agente en su clúster de EMR, tal y como se describe en Creación de un clúster de EMR que utilice el agente de Amazon CloudWatch .

Temas

- Configuración CloudWatch del agente para Amazon EMR 7.1.0
- Configuración CloudWatch del agente para Amazon EMR 7.0.0

Configuración CloudWatch del agente para Amazon EMR 7.1.0

A partir de Amazon EMR 7.1.0, puede configurar al CloudWatch agente de Amazon para que utilice métricas adicionales del sistema, añada métricas de aplicaciones y cambie el destino de las métricas mediante la API de configuración de Amazon EMR. Para obtener más información sobre cómo usar la API de configuración de EMR para configurar las aplicaciones del clúster, consulte Configuración de aplicaciones.

> **Note**
>
> 7.1.0 solo admite el tipo de reconfiguración `OVERWRITE`. Para obtener más información sobre los tipos de reconfiguración, consulte Consideraciones a la hora de reconfigurar un grupo de instancias.

Temas

- Esquema de configuración
- Ejemplos de configuraciones de métricas del sistema
- Ejemplos de configuraciones de métricas de aplicación
- Amazon Managed Service para Prometheus

Esquema de configuración

`emr-metrics` tiene las siguientes clasificaciones:

- `emr-system-metrics`: configura las métricas del sistema, como CPU, disco y memoria.
- `emr-hadoop-hdfs-datanode-metrics`— configurar las métricas JMX de Hadoop DataNode
- `emr-hadoop-hdfs-namenode-metrics`— configurar las métricas de Hadoop JMX NameNode
- `emr-hadoop-yarn-nodemanager-metrics`— configurar las métricas de Yarn JMX NodeManager
- `emr-hadoop-yarn-resourcemanager-metrics`— configurar las métricas de Yarn ResourceManager JMX
- `emr-hbase-master-metrics`— configurar las métricas de HBase Master JMX
- `emr-hbase-region-server-metrics`— configurar las métricas HBase JMX de Region Server
- `emr-hbase-rest-server-metrics`— configurar las métricas HBase JMX del servidor REST
- `emr-hbase-thrift-server-metrics`— configurar las métricas JMX de HBase Thrift Server

Las tablas siguientes describen las propiedades y configuraciones disponibles para todas y cada una de las clasificaciones.

propiedades de emr-metrics

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics_destination`	Opcional	Determina si las métricas del clúster se publican en Amazon CloudWatch o en Amazon Managed Service para Prometheus.	"CLOUDWATCH"	"CLOUDWATCH", "PROMETHEUS"	Esta propiedad no distingue entre mayúsculas y minúsculas. Por ejemplo, “Cloudwatch” es lo mismo que “CLOUDWATCH”.
`prometheus_endpoint`	Opcional	Si `metrics_destination` se establece en «PROMETHEUS», esta propiedad configura al CloudWatch agente para que envíe métricas al punto de conexión de escritura remota proporcionado por	N/A	Cualquier URL de escritura remota de Amazon Managed Service para Prometheus. El formato de URL de escritura remota es `https://aps-workspaces.<region>.com/workspaces/<workspac`	Este campo solo se obligatorio si `metrics_destination` está establecido en "PROMETHEUS". El aprovisionamiento fallará si no proporciona una clave o si el valor es una cadena vacía.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
		Amazon Managed Service para Prometheus.		`e_id> / api/v1/remote_write`	

emr-system-metrics propiedades

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas desde el CloudWatch agente.	"60"	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Puede anular esta propiedad con la propiedad `metrics_collection_interval` de grupos de métricas individuales.

emr-system-metrics configuraciones

cpu

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de la CPU que debe recopilar el agente.	Consulte Métricas predeterminadas para el CloudWatch agente con Amazon EMR	Una lista separada por comas de nombres de métricas de la CPU válidos con o sin el prefijo `cpu_`, como `usage_active` y `cpu_time_idle`. Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	Si se especifica una cadena vacía, no se publica ninguna métrica de la CPU.
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas de la CPU que	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad del `metrics_collection_interval` global solo para las

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
		debe realizar el agente.			métricas de la CPU.
`drop_original_metrics`	Opcional	Lista de métricas de la CPU para las que no se deben publicar métricas no agregadas.	No se ha publicado ninguna métrica de CPU no agregada.	Una lista separada por comas de métricas de la CPU que también se especifican en la propiedad de métricas. Una cadena vacía indica que se publican todas las métricas de la CPU.	El CloudWatch agente agrega todas las métricas por ID de clúster, ID de instancia, tipo de nodo y nombre de servicio. De forma predeterminada, el CloudWatch agente no publica las métricas por recurso de las métricas con varios recursos.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`resources`	Opcional	Determina si el agente publicará las métricas por núcleo.	"*"	"*" habilita las métricas por núcleo. "" deshabilita las métricas por núcleo.	El CloudWatch agente solo publica métricas por núcleo para las métricas de la CPU que no se descartan. `drop_original_metrics`

disk

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de disco que debe recopilar el agente.	Consulte Métricas predeterminadas para el CloudWatch agente con Amazon EMR	Una lista separada por comas de nombres de métricas de disco válidos con o sin el prefijo `disk_`, como `disk_total` y	Si se especifica una cadena vacía, no se publicará ninguna métrica de disco.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
				`used_percent` . Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas del disco que debe realizar el agente.	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad de `metrics_collection_interval` global solo para las métricas de disco.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`drop_original_metrics`	Opcional	Lista de métricas de disco para las que no se deben publicar métricas no agregadas.	No se ha publicado ninguna métrica de disco no agregada.	Una lista separada por comas de métricas de disco que también se especifican en la propiedad de métricas. Una cadena vacía indica que se publican todas las métricas del disco.	El CloudWatch agente agrega todas las métricas por ID de clúster, ID de instancia, tipo de nodo y nombre de servicio. De forma predeterminada, el CloudWatch agente no publica las métricas por recurso de las métricas con varios recursos.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`resources`	Opcional	Determina si el agente publicará per-mount-point las métricas.	"*"	"*" significa todos los puntos de montaje, "" significa que no hay puntos de montaje o se trata de una lista de puntos de montaje separados por comas. Por ejemplo, `"/,/emr"`.	El CloudWatch agente solo publica per-mount-point métricas de disco que no se descartan `drop_original_metrics` .

diskio

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de E/S de disco que debe recopilar el agente.	Consulte Métricas predeterminadas para el CloudWatch agente con Amazon EMR	Una lista separada por comas de nombres de métricas de E/S de disco válidos con o sin el prefijo`diskio_`	Si se especifica una cadena vacía, no se publicará ninguna métrica de E/S del disco.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
				como `diskio_reads` y `writes`. Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas de E/S del disco que debe realizar el agente.	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad de `metrics_collection_interval` global solo para las métricas de E/S del disco.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`drop_original_metrics`	Opcional	Lista de métricas de E/S del disco para las que no se deben publicar métricas no agregadas.	No se ha publicado ninguna métrica de E/S del disco no agregada.	Una lista separada por comas de métricas de E/S del disco que también se especifican en la propiedad de métricas. Una cadena vacía indica que se publican todas las métricas de E/S del disco.	El CloudWatch agente agrega todas las métricas por ID de clúster, ID de instancia, tipo de nodo y nombre de servicio. De forma predeterminada, el CloudWatch agente no publica las métricas por recurso de las métricas con varios recursos.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`resources`	Opcional	Determina si el agente publicará las métricas por dispositivo.	"*"	"*" significa todos los dispositivos de almacenamiento, "" significa que no hay dispositivos de almacenamiento o se trata de una lista de nombres de dispositivos separados por comas. Por ejemplo, `"nvme0n1, nvme1n1"` .	El CloudWatch agente solo publica métricas por dispositivo para las métricas de E/S de disco que no se descartan. `drop_original_metrics`

mem

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de memoria que debe	Consulte Métricas predeterminadas para	Una lista separada por comas de nombres	Si se especifica una cadena vacía, no

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
		recopilar el agente.	el CloudWatch agente con Amazon EMR	de métricas de memoria válidos con o sin el prefijo `mem_`, como `mem_available` y `available_percent`. Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	se publica ninguna métrica de memoria.
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas de memoria que debe realizar el agente.	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad de `metrics_collection_interval` global solo para las métricas de memoria.

net

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de red que debe recopilar el agente.	Consulte Métricas predeterminadas para el CloudWatch agente con Amazon EMR	Una lista separada por comas de nombres de métricas de red válidos con o sin el prefijo`net_`, como `net_packets_sent` y `packets_recv` . Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	Si se especifica una cadena vacía, no se publica ninguna métrica de red.
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas de red que debe	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad de `metrics_collection_interval` global solo para las métricas de red.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
		realizar el agente.			
`drop_original_metrics`	Opcional	Lista de métricas de red para las que no se deben publicar métricas no agregadas.	No se ha publicado ninguna métrica de red no agregada.	Una lista separada por comas de métricas de red que también se especifican en la propiedad de métricas. Una cadena vacía indica que se publican todas las métricas de red.	El CloudWatch agente agrega todas las métricas por ID de clúster, ID de instancia, tipo de nodo y nombre de servicio. De forma predeterminada, el CloudWatch agente no publica las métricas por recurso de las métricas con varios recursos.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`resources`	Opcional	Determina si el agente publicará las métricas por interfaz.	"*"	"*" se refiere a todas las interfaces de red, "" significa que no hay interfaces de red o se trata de una lista de nombres de interfaces separados por comas. Por ejemplo, `"eth0,eth1"` .	El CloudWatch agente solo publica métricas por interfaz para las métricas de red que no se descartan . `drop_original_metrics`

netstat

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de estadísticas de red que debe recopilar el agente.	Consulte Métricas predeterminadas para el CloudWatch agente con Amazon EMR	Una lista separada por comas de nombres de métricas de memoria válidos con o sin el prefijo `netstat_`,	Si se especifica una cadena vacía, no se publica ninguna métrica estadística de red.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
				como `tcp_listen` y `netstat_udp_socket`. Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas estadísticas de red que debe realizar el agente.	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad de `metrics_collection_interval` global solo para las métricas estadísticas de red.

processes

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de proceso que debe recopilar el agente.	Consulte Métricas predeterminadas para el CloudWatch agente con Amazon EMR	Una lista separada por comas de nombres de métricas de memoria válidos con o sin el prefijo `processes_` , como `processes_running` y `total`. Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	Si se especifica una cadena vacía, no se publica ninguna métrica del proceso.
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas del proceso del sistema que	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad de `metrics_collection_interval` global solo para las métricas de

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
		debe realizar el agente.			los procesos del sistema.

swap

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics`	Opcional	La lista de métricas de intercambio que debe recopilar el agente.	Consulte Métricas predeterminadas para el CloudWatch agente con Amazon EMR	Una lista separada por comas de nombres de métricas de memoria válidos con o sin el prefijo `swap_`, como `swap_free` y `used_percent` . Consulte Métricas que recopila el CloudWatch agente para ver métricas válidas.	Si se especifica una cadena vacía, no se publica ninguna métrica de intercambio.

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	Notas
`metrics_collection_interval`	Opcional	La frecuencia en segundos acerca de la recopilación y publicación de las métricas de intercambio que debe realizar el agente.	El valor del `metrics_collection_interval` global.	Una cadena que especifica el número de segundos. Solo acepta números enteros.	Este valor anula la propiedad de `metrics_collection_interval` global solo para las métricas de intercambio.

emr-hadoop-hdfs-datanode-Propiedades de métricas

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
<custom_bean_name>	Opcional	N/A	La fuente de la MBean que el CloudWatch agente debe recopilar las métricas, comoHadoop:service=DataNode,name=DataNodeActivity . Puede	Una cadena que contiene la lista delimitada por comas de métricas que están asociadas al. MBean Por ejemplo, `BlocksCached,BlocksRead` .	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
			encontrar ejemplos de MBean nombres y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.		
`otel.metric.export.interval`	Opcional	La frecuencia en milisegundos acerca de la recopilación de las métricas de Hadoop DataNode .	“60 000”	Una cadena que especifica el número de milisegundos. Solo acepta números enteros.	

emr-hadoop-hdfs-namenode-propiedades de métricas

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
`<custom_bean_name>`	Opcional	N/A	La fuente de la MBean que el	Una cadena que contiene la lista	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
			CloudWatch agente debe recopilar las métricas, como`Hadoop:service=NameNode,name=FSNamesystem` . Puede encontrar ejemplos de MBean nombres y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.	delimitada por comas de métricas que están asociadas al. MBean Por ejemplo, `BlockCapacity,CapacityUsedGB` .	
`otel.metric.export.interval`	Opcional	La frecuencia en milisegundos acerca de la recopilación de las métricas de Hadoop NameNode .	“60 000”	Una cadena que especifica el número de milisegundos. Solo acepta números enteros.	

emr-hadoop-yarn-nodemanager-propiedades de métricas

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
<custom_bean_name>	Opcional	N/A	La fuente de la MBean que el CloudWatch agente debe recopilar las métricas, como`Hadoop:service=NodeManager,name=NodeManagerMetrics` . Puede encontrar ejemplos de MBean nombres y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.	Una cadena que contiene la lista delimitada por comas de métricas que están asociadas al. MBean Por ejemplo, `MaxCapacity,AllocatedGB` .	
`otel.metric.export`	Opcional	La frecuencia en milisegundos acerca	“60 000”	Una cadena que especifica el número	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
`.interval`		de la recopilación de las métricas de Hadoop YARN NodeManager .		de milisegundos. Solo acepta números enteros.	

emr-hadoop-yarn-resourcemanager-propiedades de métricas

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
<custom_bean_name>	Opcional	N/A	La fuente de la MBean que el CloudWatch agente debe recopilar las métricas, como`Hadoop:service=ResourceManager,name=PartitionQueueMetrics` . Puede encontrar ejemplos de MBean	Una cadena que contiene la lista delimitada por comas de métricas que están asociadas al. MBean Por ejemplo, `MaxCapacity,MaxCapacityVCores` .

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
			nombres y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.		
`otel.metric.export.interval`	Opcional	La frecuencia en milisegundos acerca de la recopilación de las métricas de Hadoop YARN ResourceManager .	“60 000”	Una cadena que especifica el número de milisegundos. Solo acepta números enteros.	

emr-hbase-master-metrics propiedades

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
<custom_bean_name>	Opcional	N/A	La fuente de la MBean que el CloudWatch	Una cadena que contiene la lista delimitada	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
			agente debe recopilar las métricas, como. `Hadoop:service=HBase,name=Master,sub=AssignmentManager` Puede encontrar ejemplos de MBean nombres y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.	por comas de métricas que están asociadas al. MBean Por ejemplo, `AssignFailedCount, AssignSubmittedCount` .	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
`otel.metric.export.interval`	Opcional	La frecuencia en milisegundos acerca de la recopilación de las métricas HBase maestras.	“60 000”	Una cadena que especifica el número de milisegundos. Solo acepta números enteros.	

emr-hbase-region-server-propiedades de métricas

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
<custom_bean_name>	Opcional	N/A	La fuente de la MBean que el CloudWatch agente debe recopilar las métricas, comoHadoop:service=HBase,name=RegionServer,sub=IPC . Puede encontrar ejemplos de MBean	Una cadena que contiene la lista delimitada por comas de métricas que están asociadas al. MBean Por ejemplo, numActiveHandler,numActivePriorityHandler .	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
			nombres y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.		
`otel.metric.export.interval`	Opcional	La frecuencia en milisegundos acerca de la recopilación de las métricas de HBase Region Server.	“60 000”	Una cadena que especifica el número de milisegundos. Solo acepta números enteros.	

emr-hbase-rest-server-propiedades de métricas

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
<custom_bean_name>	Opcional	N/A	La fuente de la MBean que el CloudWatch agente debe	Una cadena que contiene la lista delimitada por comas	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
			recopilar las métricas, como`Hadoop:service=HBase,name=REST` . Puede encontrar ejemplos de MBean nombres y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.	de métricas que están asociadas al. MBean Por ejemplo, `successfulPut,successfulScanCount` .	
`otel.metric.export.interval`	Opcional	La frecuencia en milisegundos acerca de la recopilación de las métricas de HBase Rest Server.	“60 000”	Una cadena que especifica el número de milisegundos. Solo acepta números enteros.	

emr-hbase-thrift-server-propiedades de métricas

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
<custom_bean_name>	Opcional	N/A	La fuente de la MBean que el CloudWatch agente debe recopilar las métricas, como`Hadoop:service=HBase,name=Thrift,sub=ThriftOne` . Puede encontrar MBean nombres de ejemplo y sus métricas correspondientes en los archivos JMX YAML de ejemplo para Amazon EMR versión 7.0.	Una cadena que contiene la lista delimitada por comas de métricas que están asociadas al. MBean Por ejemplo, `BatchGet_max,BatchGet_mean` .	
`otel.metric.export.interval`	Opcional	La frecuencia en milisegundos acerca de la recopilac	“60 000”	Una cadena que especifica el número de milisegundos. Solo	

Propiedad	Obligatorio	Descripción	Valor predeterminado	Valores posibles	
		ión de las métricas de HBase Thrift Server.		acepta números enteros.	

Ejemplos de configuraciones de métricas del sistema

En el siguiente ejemplo se muestra cómo configurar el CloudWatch agente para detener la exportación de todas las métricas del sistema.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

En el siguiente ejemplo configura al CloudWatch agente para exportar las métricas del sistema. Hacerlo es una forma rápida de restablecer el agente para que solo exporte las métricas predeterminadas del sistema si ya ha reconfigurado las métricas del sistema al menos una vez. Este restablecimiento también elimina cualquier métrica de la aplicación que se haya reconfigurado anteriormente.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": []
  }
```

```
]
```

El siguiente ejemplo configura el clúster para exportar las métricas `cpu`, `mem` y `disk`.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {
          "metrics_collection_interval": "20"
        },
        "Configurations": [
          {
            "Classification": "cpu",
            "Properties": {
              "metrics": "cpu_usage_guest,cpu_usage_idle",
              "metrics_collection_interval": "30",
              "drop_original_metrics": "cpu_usage_guest"
            }
          },
          {
            "Classification": "mem",
            "Properties": {
              "metrics": "mem_active"
            }
          },
           {
            "Classification": "disk",
            "Properties": {
              "metrics": "disk_used_percent",
              "resources": "/,/mnt",
              "drop_original_metrics": ""
            }
          }
        ]
      }
    ]
  }
]
```

La configuración del ejemplo anterior tiene las propiedades siguientes:

- Cada 30 segundos, el agente recopila la `cpu_guest` métrica de todos CPUs. Puede encontrar la métrica agregada en el espacio de CloudWatch nombres`CWAgent > cluster.id, instance.id, node.type, service.name`.
- Cada 30 segundos, el agente recopila la `cpu_idle` métrica de todos CPUs. Puede encontrar la métrica agregada en el espacio de CloudWatch nombres`CWAgent > cluster.id, instance.id, node.type, service.name`. El agente también recopila las métricas por CPU. Puede encontrarlas en el mismo espacio de nombres. El agente recopila esta métrica porque la propiedad `drop_original_metrics` no contiene la `cpu_idle`, por lo que el agente no ignora la métrica.
- Cada 20 segundos, el agente recopila la métrica `mem_active`. Puedes encontrar la métrica agregada en el espacio de nombres. CloudWatch `CWAgent > cluster.id, instance.id, node.type, service.name`
- Cada 20 segundos, el agente recopila las métricas `disk_used_percent` de los montajes del disco `/` y `/mnt`. Puedes encontrar las métricas agregadas en el espacio de nombres. CloudWatch `CWAgent > cluster.id, instance.id, node.type, service.name` El agente también recopila las métricas por montaje. Puede encontrarlas en el mismo espacio de nombres. El agente recopila esta métrica porque la propiedad `drop_original_metrics` no contiene la `disk_used_percent`, por lo que el agente no ignora la métrica.

Ejemplos de configuraciones de métricas de aplicación

En el siguiente ejemplo configura al CloudWatch agente para detener la exportación de las métricas para el servicio Hadoop Namenode.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
```

```
]
```

El siguiente ejemplo configura un clúster para exportar las métricas de las aplicaciones de Hadoop.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {
          "Hadoop:service=NameNode,name=FSNamesystem": "BlockCapacity,CapacityUsedGB",
          "otel.metric.export.interval": "20000"
        },
        "Configurations": []
      },
       {
        "Classification": "emr-hadoop-hdfs-datanode-metrics",
        "Properties": {
          "Hadoop:service=DataNode,name=JvmMetrics": "MemNonHeapUsedM",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      },
       {
        "Classification": "emr-hadoop-yarn-resourcemanager-metrics",
        "Properties": {
          "Hadoop:service=ResourceManager,name=CapacitySchedulerMetrics":
 "AllocateNumOps,NodeUpdateNumOps"
        },
        "Configurations": []
      }
    ]
  }
]
```

El ejemplo anterior tiene las propiedades siguientes:

- Cada 20 segundos, el agente recopila las métricas `BlockCapacity` y `CapacityUsedGB` de las instancias que ejecutan el servicio Hadoop Namenode.

- Cada 30 segundos, el agente recopila las métricas `MemNonHeapUsedM` de las instancias que ejecutan el servicio Hadoop Datanode.
- Cada 30 segundos, el agente recopila las `NodeUpdateNumOps` métricas `AllocateNumOps` y métricas de las instancias que ejecutan el Hadoop YARN. ResourceManaager

Amazon Managed Service para Prometheus

En el siguiente ejemplo se muestra cómo configurar el CloudWatch agente para exportar métricas a Amazon Managed Service para Prometheus.

Si actualmente exporta métricas a Amazon Managed Service para Prometheus y desea reconfigurar las métricas del clúster y seguir exportándolas a Amazon Managed Service para Prometheus, debe incluir las propiedades `metrics_destination` y `prometheus_endpoint`.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "http://amp-workspace/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

Para usar el CloudWatch agente al que exportar métricas CloudWatch, utilice el siguiente ejemplo.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

> **Note**
>
> El CloudWatch agente tiene un exportador de Prometheus que cambia el nombre de algunos atributos. Para las etiquetas de métricas predeterminadas, Amazon Managed Service para Prometheus utiliza caracteres de subrayado en lugar de los puntos que utiliza Amazon. CloudWatch Si utiliza Amazon Managed Grafana para visualizar las métricas predeterminadas en Amazon Managed Service para Prometheus, las etiquetas aparecen como `cluster_id`, `instance_id`, `node_type` y `service_name`.

Configuración CloudWatch del agente para Amazon EMR 7.0.0

Puede configurar el CloudWatch agente de Amazon para que utilice métricas del sistema adicionales, diferentes de las que proporciona la configuración predeterminada del CloudWatch agente. La configuración de la versión 7.0.0 requiere el uso de acciones de arranque, de las que se ofrecen ejemplos en las siguientes secciones. En una próxima versión, Amazon EMR ofrecerá opciones de configuración adicionales a través de la API de Amazon EMR.

Temas

- Configuración de métricas de sistema adicionales con Amazon EMR 7.0.0
- Configuración de las métricas de las aplicaciones con Amazon EMR 7.0.0
- Configuración de Amazon Managed Service para Prometheus como almacenamiento en la nube para métricas con Amazon EMR 7.0.0

Configuración de métricas de sistema adicionales con Amazon EMR 7.0.0

Realice los siguientes pasos para configurar el agente para que utilice un conjunto diferente de métricas del sistema en Amazon EMR 7.0.0:

1. Cree o elija un bucket en la cuenta de Amazon S3 en la que desea almacenar los archivos de configuración que especifican las métricas del CloudWatch agente.
2. Cree el archivo de configuración del `emr-amazon-cloudwatch-agent.json` con las métricas especificadas que prefiera. Para ello, utilice uno de los métodos explicados en Crear el archivo de configuración del CloudWatch agente. Para obtener más información acerca de la estructura del archivo de configuración del CloudWatch agente, consulte Crear el archivo de configuración del CloudWatch agente de forma manual en la Guía del CloudWatch usuario de Amazon.

3. A continuación, vaya al aws-emr-utilities repositorio en GitHub y descargue los siguientes scripts de métricas del sistema:

   - **`install_system_metrics_launcher.sh`**: un script que se descarga y luego ejecuta `install_system_metrics.sh` en segundo plano para que el nodo pueda terminar de arrancar.
   - **`install_system_metrics.sh`**: un script que espera a que la instancia en la que se ejecuta termine de arrancar y, a continuación, descarga y aplica la configuración en el archivo JSON.
4. Abra cada archivo SH y reemplace *`amzn-s3-demo-bucket`* por el nombre del bucket del paso 1.
5. Cargue el archivo JSON y los dos archivos SH en el bucket de S3.
6. Ahora puede ir a la consola de Amazon EMR y crear un clúster nuevo con el CloudWatch agente. En EMR activado, EC2 en el panel de navegación de la izquierda, seleccione Clústeres y, a continuación, Crear clúster.
7. En la sección Nombre y aplicaciones, elija la versión de Amazon EMR 7.0.0 o posterior.
8. En Paquete de aplicaciones, seleccione el paquete o las aplicaciones que desea instalar en el clúster e incluya el CloudWatch agente de Amazon en la selección.
9. En la sección Acciones de arranque, seleccione Agregar.

   - En Nombre, ingrese **`install_system_metrics_launcher.sh`**.
   - Para la Ubicación del script, ingrese **`s3://amzn-s3-demo-bucket/install_system_metrics_launcher.sh`**. Reemplace *`amzn-s3-demo-bucket`* por la ruta al bucket de S3.
   - Deje el bloque Argumentos vacío.
10. Seleccione Agregar acción de arranque.
11. Continúe con la creación del clúster para satisfacer las necesidades de la carga de trabajo.

Cuando se lanza el clúster, el CloudWatch agente publica las métricas del sistema que especificó en el archivo de configuración CloudWatch.

Configuración de las métricas de las aplicaciones con Amazon EMR 7.0.0

Puede configurar el CloudWatch agente de Amazon para que publique métricas de aplicaciones para HDFS y YARN además de las métricas del sistema. Realice los pasos siguientes para configurar el agente para que publique métricas de aplicaciones:

1. Cree o elija un bucket en la cuenta de Amazon S3 en la que desea almacenar los archivos de configuración que especifican las métricas del CloudWatch agente.
2. A continuación, vaya al aws-emr-utilities repositorio en GitHub y descargue los siguientes scripts:
   - **`install_app_metrics_launcher.sh`**: un script que se descarga y luego ejecuta `install_app_metrics.sh` en segundo plano para que el nodo pueda terminar de arrancar.
   - **`install_app_metrics.sh`**: un script que espera a que la instancia en la que se ejecuta termine de arrancar y, a continuación, descarga y aplica la configuración en los archivos YAML que se descargarán en el paso siguiente.
3. Abra cada archivo y reemplace *`amzn-s3-demo-bucket`* por el nombre del bucket del paso 1.
4. A continuación, descargue los siguientes archivos de mapeo YAML. Para obtener información sobre cómo están estructurados estos archivos YAML, consulta `javaagent`en el OpenTelemetry Instrumentation for Java GitHub repositorio.
   - **`datanode-metrics.yaml`**: la configuración de las métricas de Hadoop DataNode.
   - **`namenode-metrics.yaml`**: la configuración de las métricas de Hadoop NameNode.
   - **`nodemanager-metrics.yaml`**: la configuración de las métricas de Yarn NodeManager.
   - **`resourcemanager-metrics.yaml`**: la configuración de las métricas de Yarn ResourceManager.
5. Cargue los dos archivos SH y los cuatro archivos YAML en el bucket de S3.
6. Ahora puede ir a la consola de Amazon EMR y crear un clúster nuevo con el CloudWatch agente. En EMR activado, EC2 en el panel de navegación de la izquierda, seleccione Clústeres y, a continuación, Crear clúster.
7. En la sección Nombre y aplicaciones, elija la versión de Amazon EMR 7.0.0 o posterior.
8. En Paquete de aplicaciones, seleccione el paquete o el grupo personalizado de aplicaciones que desee instalar en el clúster e incluya el CloudWatch agente en la selección.
9. En la sección Acciones de arranque, seleccione Agregar.
   - En Nombre, ingrese **`install_app_metrics_launcher.sh`**.
   - Para la Ubicación del script, ingrese **`s3://amzn-s3-demo-bucket/install_app_metrics_launcher.sh`**. Reemplace *`amzn-s3-demo-bucket`* por la ruta al bucket de S3.
   - Deje el bloque Argumentos vacío.
10. Seleccione Agregar acción de arranque.

11. Continúe con la creación del clúster para satisfacer las necesidades de la carga de trabajo.

Cuando se lanza el clúster, el CloudWatch agente publica las métricas de la aplicación que especificó junto con las métricas del sistema CloudWatch.

Configuración de Amazon Managed Service para Prometheus como almacenamiento en la nube para métricas con Amazon EMR 7.0.0

Puede configurar el CloudWatch agente de Amazon para que publique métricas en Amazon Managed Service para Prometheus en lugar de. CloudWatch

> **Note**
>
> Puede publicar métricas del CloudWatch agente de Amazon en Amazon Managed Service para Prometheus o en CloudWatch Amazon, pero no puede publicar las métricas en ambos servicios en un mismo clúster.

Para configurar el agente para que publique métricas en Amazon Managed Service para Prometheus, debe agregar `aps:RemoteWrite` AWS Identity and Access Management el permiso (IAM) IAM al perfil de EC2 instancia de Amazon para Amazon EMR. El siguiente ejemplo de política contiene los permisos necesarios:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "aps:RemoteWrite",
      "Resource": "*"
    }
  ]
}
```

Uso del CloudWatch agente en un clúster de EMR para que publique métricas en Amazon Managed Service para Prometheus

Cuando la política de servicio tenga los permisos correctos, realice los siguientes pasos para lanzar un clúster que utilice el CloudWatch agente para publicar métricas en Amazon Managed Service para Prometheus.

1. Utilice la AWS Management Console o la AWS CLI para crear un espacio de trabajo de Amazon Managed Service para Prometheus. Para obtener más información, consulte Crear un espacio de trabajo en la Guía del usuario de Amazon Managed Service para Prometheus.
2. Cree o elija un bucket en la cuenta de Amazon S3 en la que desea almacenar los archivos de lanzamiento que especifican Amazon Managed Service para Prometheus como almacenamiento en la nube.
3. A continuación, vaya al aws-emr-utilities repositorio en GitHub y descargue los siguientes scripts:
   - **`add_prometheus_endpoint_launcher.sh`**: un script que se descarga y luego ejecuta `add_prometheus_endpoint.sh` en segundo plano para que el nodo pueda terminar de arrancar.
   - **`add_prometheus_endpoint.sh`**— un script que espera a que termine el arranque de la instancia en la que se ejecuta y, a continuación, configura el CloudWatch agente a fin de publicar en el punto de conexión de Amazon Managed Service para Prometheus que se proporciona como argumento al lanzar el clúster.
4. Abra cada archivo y reemplace *`amzn-s3-demo-bucket`* por el nombre del bucket del paso 2.
5. Utilice AWS CLI para crear un clúster de EMR con la acción de `add_prometheus_endpoint_launcher.sh` arranque. En el siguiente comando, reemplace *`amzn-s3-demo-bucket`* por el bucket que contiene la acción de arranque y reemplace *`managedpro-remote-write-workspace-url`* por el punto de conexión de escritura remoto del espacio de trabajo de Amazon Managed Service para Prometheus. Asegúrese de especificar una etiqueta de versión de Amazon EMR igual a `emr-7.0.0` o posterior.

```
aws emr create-cluster --name managedpro-cluster \
    --release-label emr-7.0.0 \
    --applications Name=Hadoop Name=AmazonCloudWatchAgent \
    --ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
    --instance-count 3 --use-default-roles
```

```
  --bootstrap-actions Name='Add Prometheus Endpoint',Path=s3://amzn-s3-demo-
bucket/add_prometheus_endpoint_launcher.sh,Args='managedpro-remote-write-workspace-
url'
```

Cuando se lanza el clúster, el CloudWatch agente publica las métricas que recopila en Amazon Managed Service para Prometheus.

Uso de Amazon Managed Service para Prometheus como origen de datos para Amazon Managed Grafana

Una vez que Amazon EMR haya publicado las métricas del clúster en Amazon Managed Service para Prometheus, puede realizar los siguientes pasos para visualizar las métricas con Amazon Managed Grafana:

1. Utilice AWS Management Console para crear un espacio de trabajo y un usuario de Amazon Managed Grafana con los permisos adecuados. Para obtener más información, consulte Crear un espacio de trabajo en la Guía del usuario de Amazon Managed Grafana.
2. Agregue Amazon Managed Service para Prometheus como origen de datos para Amazon Managed Grafana. Para obtener más información, consulte Uso de la configuración de la fuente de datos AWS para agregar Amazon Managed Service para Prometheus como origen de datos en la Guía del usuario de Amazon Managed Grafana.

> **Note**
>
> El CloudWatch agente tiene un exportador de Prometheus que cambia el nombre de algunos atributos. Para las etiquetas de métricas predeterminadas, Amazon Managed Service para Prometheus utiliza caracteres de subrayado en lugar de los puntos que utiliza Amazon. CloudWatch Por lo tanto, si utiliza Amazon Managed Grafana para visualizar las métricas predeterminadas en Amazon Managed Service para Prometheus, las etiquetas aparecen como `jobflow_id`, `instance_id` y `service_name`.
> Además, cualquier métrica de la aplicación que el CloudWatch agente publique en Amazon Managed Service para Prometheus utiliza la `job` etiqueta. `service_name` Sin embargo, las métricas del sistema siguen utilizando la etiqueta `service_name`.

Consideraciones y limitaciones

- El CloudWatch agente de Amazon nativo está disponible para los clústeres que cree con las versiones 7.0.0 y posteriores de Amazon EMR.
- La configuración del CloudWatch agente en Amazon EMR 7.0.0 requiere el uso de acciones de arranque. En una próxima versión, Amazon EMR ofrecerá opciones de configuración adicionales a través de la API de Amazon EMR.
- No puede instalar el CloudWatch agente de Amazon EMR si ya lo implementó mediante otro método, como CloudWatch Agent (SSM AWS Systems Manager Agent), en la región en la que creó el clúster. Si lo hace, el clúster se cerrará con errores.
- La CloudWatch GetMetricData API admite hasta 500 métricas por solicitud. Si su clúster de Amazon EMR contiene más de 250 nodos en un grupo de instancias o una flota, los gráficos correspondientes del panel CloudWatch integrado en EMR aparecerán en blanco con el error Demasiadas métricas, ya que estas métricas requieren dos puntos de datos por métrica en el panel de información general del clúster. Sin embargo, si filtra los paneles principales o del grupo de instancias de tareas (o flota), podrá ver los gráficos de hasta 500 nodos por grupo de instancias o flota. Esto se debe a que no requieren dos puntos de datos por métrica. Si hay más de 500 nodos por grupo o flota de instancias, también se produce el error de demasiadas métricas en las métricas de estos paneles.

CloudWatch historial de lanzamientos del agente

En la siguiente tabla se muestra la versión AmazonCloudWatchAgent incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

AmazonCloudWatchAgent información sobre la versión

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
emr7.9.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies,

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
		emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr7.8.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
emr-7.7.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.6.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
emr-7.5.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7.4.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
emr-7.3.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7.2.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	AmazonCloudWatchAgent Versión	Componentes instalados con AmazonCloudWatchAgent
emr-7.1.0	1.300032,2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.0.0	1.300031,01	adot-java-agent, emr-amazon-cloudwatch-agent

Delta Lake

Delta Lake es un marco de capas de almacenamiento para arquitecturas de lake house que se suelen basar en Amazon S3. Con las versiones 6.9.0 y posteriores de Amazon EMR, puede utilizar Apache Spark 3.x en clústeres de Amazon EMR con tablas de Delta Lake. Para obtener más información sobre los lake houses con Delta Lake, consulte https://delta.io/.

En la siguiente tabla, se muestra la versión de Delta incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Delta.

Para ver la versión de los componentes instalados con Delta en esta versión, consulte Versiones de componentes de la versión 7.9.0.

Información sobre la versión Delta para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Delta	Componentes instalados con Delta
emr-7.9.0	Delta 3.3.0-amzn-1	Not available.

En la tabla siguiente, se muestra la versión de Delta incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Delta.

Para ver la versión de los componentes instalados con Delta en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Delta para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Delta	Componentes instalados con Delta
emr-6.15.0	Delta 2.4.0	Not available.

Introducción a Delta Lake

Delta Lake es un proyecto de código abierto que ayuda a implementar arquitecturas de lago de datos modernas que se suelen basar en Amazon S3. Data Lake ofrece las siguientes capacidades:

- Transacciones atómicas, coherentes, aisladas y duraderas (ACID) en Spark. Los lectores ven una vista coherente de la tabla durante un trabajo de Spark.
- Gestión escalable de metadatos con procesamiento distribuido por Spark.
- Combina los casos de uso de streaming y lotes en la misma tabla Delta.
- Aplicación automática del esquema para evitar registros incorrectos durante la ingesta de datos.
- Viaje en el tiempo con el control de versiones de datos.
- Admite operaciones de combinación, actualización y eliminación para casos de uso complejos, como la captura de datos de cambio (CDC), upserts de streaming y mucho más.

Uso de un clúster con Delta Lake instalado

Temas

- Uso de un clúster de Delta Lake con Flink
- Utilice un clúster de Delta Lake con Trino
- Utilice un clúster de Delta Lake con Spark
- Uso de un clúster de Delta Lake con Spark y AWS Glue

Uso de un clúster de Delta Lake con Flink

Con la versión 6.11 y posteriores de Amazon EMR, puede usar Delta Lake con su clúster de Flink. En los siguientes ejemplos, se utiliza AWS CLI para trabajar con Delta Lake en un clúster Flink de Amazon EMR.

> **Note**
>
> Amazon EMR admite la DataStream API de Flink cuando utiliza Delta Lake con un clúster de Flink.

Creación de un clúster de Delta Lake

1. Cree un archivo, `delta_configurations.json`, con el siguiente contenido:

```
[{"Classification":"delta-defaults",
```

```
    "Properties":{"delta.enabled":"true"}}]
```

2. Cree un clúster con la siguiente configuración. Reemplace `example Amazon S3 bucket path` y `subnet ID` por sus valores.

```
aws emr create-cluster
--release-label emr-6.11.0
--applications Name=Flink
--configurations file://delta_configurations.json
--region us-east-1  --name My_Spark_Delta_Cluster
--log-uri  s3://amzn-s3-demo-bucket/
--instance-type m5.xlarge
--instance-count 3
--service-role EMR_DefaultRole_V2
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Inicialización de una sesión de yarn de Flink

Para iniciar una sesión de Flink Yarn, ejecute el siguiente comando:

```
flink-yarn-session -d
```

Creación de un trabajo de Flink con Delta Lake

En los siguientes ejemplos, se muestra cómo usar sbt o Maven para crear su trabajo de Flink con Delta Lake.

sbt

sbt es una herramienta de compilación para Scala que puede usar con poca o ninguna configuración cuando tiene proyectos pequeños.

```
libraryDependencies ++= Seq(
  "io.delta" %% "delta-flink" % deltaConnectorsVersion % "provided",
  "io.delta" %% "delta-standalone" % deltaConnectorsVersion % "provided",
  "org.apache.flink" %% "flink-clients" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-parquet" % flinkVersion % "provided",
  "org.apache.hadoop" % "hadoop-client" % hadoopVersion % "provided",
  "org.apache.flink" % "flink-table-common" % flinkVersion % "provided",
```

```
  "org.apache.flink" %% "flink-table-runtime" % flinkVersion % "provided")
```

Maven

Maven es una herramienta de automatización de compilaciones de código abierto de Apache Software Foundation. Con Maven, puede crear, publicar e implementar un trabajo de Flink con Delta Lake en Amazon EMR.

```
<project>
<properties>
    <scala.main.version>2.12</scala.main.version>
    <delta-connectors-version>0.6.0</delta-connectors-version>
    <flink-version>1.16.1</flink-version>
    <hadoop-version>3.1.0</hadoop-version>
</properties>

<dependencies>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-flink</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-standalone_$scala-main-version</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-clients</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-parquet</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.hadoop</groupId>
```

```
        <artifactId>hadoop-client</artifactId>
        <version>$hadoop-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-common</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-runtime</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
</dependencies>
```

Escritura en una tabla Delta con la API DataStream de Flink

Utilice el siguiente ejemplo para crear un objeto DeltaSink para escribir en la tabla con un `deltaTablePath`:

```
public static DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath,
        RowType rowType) {
    Configuration configuration = new Configuration();
    DeltaSink<RowData> deltaSink = DeltaSink
            .forRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration,
                    rowType)
            .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Lectura de una tabla Delta con la API Flink Datastream

Utilice el siguiente ejemplo para crear un acotado DeltaSource que se lea en la tabla con un `deltaTablePath`:

```
public static DataStream<RowData> createBoundedDeltaSourceAllColumns(
        StreamExecutionEnvironment env,
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    DeltaSource<RowData> deltaSource = DeltaSource
            .forBoundedRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration)
            .build();

    return env.fromSource(deltaSource, WatermarkStrategy.noWatermarks(), "delta-
source");
}
```

Creación de receptores con soporte para múltiples clústeres para la versión independiente de Delta Lake

Utilice el siguiente ejemplo para crear una tabla en la que DeltaSink escribir `deltaTablePath` y que sea compatible con varios clústeres:

```
public DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    configuration.set("spark.delta.logStore.s3.impl",
 "io.delta.storage.S3DynamoDBLogStore");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.tableName",
 "delta_log");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.region", "us-
east-1");

    DeltaSink<RowData> deltaSink = DeltaSink
        .forRowData(
            new Path(deltaTablePath),
            configuration,
            rowType)
        .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Ejecución del trabajo de Flink

Use el comando siguiente para ejecutar el trabajo:

```
flink run FlinkJob.jar
```

Utilice un clúster de Delta Lake con Trino

Con Amazon EMR versión 6.9.0 y versiones posteriores, puede usar Delta Lake con su clúster de Trino.

En este tutorial, utilizaremos el clúster Trino AWS CLI para trabajar con Delta Lake en Amazon EMR.

Creación de un clúster de Delta Lake

1. Cree un archivo, `delta_configurations.json` y establezca valores para el catálogo que elija. Por ejemplo, si quiere usar el metaalmacén de Hive como catálogo, su archivo debe tener el siguiente contenido:

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore.uri":"thrift://localhost:9083"}}]
   ```

   Si quieres usar el catálogo de AWS Glue como tienda, tu archivo debe tener el siguiente contenido:

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore":"glue"}}]
   ```

2. Cree un clúster con la siguiente configuración y sustituya **`example Amazon S3 bucket path`** y **`subnet ID`** por sus propios valores.

   ```
   aws emr create-cluster
       --release-label emr-6.9.0
       --applications Name=Trino
       --configurations file://delta_configurations.json
   ```

```
    --region us-east-1  --name My_Spark_Delta_Cluster
    --log-uri  s3://amzn-s3-demo-bucket/
    --instance-type m5.xlarge
    --instance-count 2
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Inicialización de una sesión de Trino para Delta Lake

Para iniciar una sesión de Trino, ejecute el siguiente comando

```
trino-cli --catalog delta
```

Escritura en una tabla de Delta Lake

Cree la siguiente tabla y escriba en ella con los siguientes comandos de SQL:

```
SHOW SCHEMAS;

CREATE TABLE default.delta_table (id  int, data varchar, category varchar) WITH
( location =  's3://amzn-s3-demo-bucket/<prefix>');

INSERT INTO default.delta_table VALUES  (1,'a','c1'), (2,'b','c2'), (3,'c','c3');
```

Lectura de una tabla Delta Lake

Lea desde la tabla con el siguiente comando SQL:

```
SELECT * from default.delta_table;
```

Utilice un clúster de Delta Lake con Spark

A partir de la versión 6.9.0 de Amazon EMR, puede usar Delta Lake con su clúster de Spark sin necesidad de incluir acciones de arranque. Para los lanzamientos 6.8.0 y anteriores de Amazon EMR, puede utilizar acciones de arranque para preinstalar todas las dependencias necesarias.

En los siguientes ejemplos, se utiliza AWS CLI para trabajar con Delta Lake en un clúster de Amazon EMR Spark.

Para usar Delta Lake en Amazon EMR con AWS Command Line Interface, primero cree un clúster. Para obtener información sobre cómo especificar la clasificación de Delta Lake con AWS Command Line Interface, consulte Suministrar una configuración mediante la AWS Command Line Interface al crear un clúster o Proporcionar una configuración con el SDK de Java al crear un clúster.

1. Cree un archivo, `configurations.json`, con el siguiente contenido:

```
[{"Classification":"delta-defaults",  "Properties":{"delta.enabled":"true"} }]
```

2. Cree un clúster con la siguiente configuración y sustituya el ejemplo de Amazon S3 **`bucket path`** y **`subnet ID`** por el suyo propio.

```
aws emr create-cluster
     --release-label  emr-6.9.0
     --applications Name=Spark
     --configurations file://delta_configurations.json
     --region us-east-1
     --name My_Spark_Delta_Cluster
     --log-uri  s3://amzn-s3-demo-bucket/
     --instance-type m5.xlarge
     --instance-count 2
     --service-role EMR_DefaultRole_V2
     --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

   Como alternativa, puede crear un clúster de Amazon EMR y una aplicación de Spark con los siguientes archivos como dependencias de JAR en un trabajo de Spark:

```
/usr/share/aws/delta/lib/delta-core.jar,
/usr/share/aws/delta/lib/delta-storage.jar,
/usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar
```

> **Note**
>
> Si utiliza Amazon EMR versión 7.0.0 o superior, `/usr/share/aws/delta/lib/delta-spark.jar` utilícela en lugar de. `/usr/share/aws/delta/lib/delta-core.jar`

Para obtener más información, consulte Envío de aplicaciones.

Para incluir una dependencia jar en un trabajo de Spark, agregue las siguientes propiedades de configuración a la aplicación Spark:

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-core.jar,
     /usr/share/aws/delta/lib/delta-storage.jar,
     /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Para obtener más información sobre las dependencias de trabajos de Spark, consulte Dependency Management.

Si utiliza Amazon EMR versión 7.0.0 o superior, añada la `/usr/share/aws/delta/lib/delta-spark.jar` configuración en su lugar.

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-spark.jar,
     /usr/share/aws/delta/lib/delta-storage.jar,
     /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Inicio de sesión de Spark para Delta Lake

En los siguientes ejemplos se muestra cómo iniciar el intérprete de comandos interactivo de Spark, utilizar el envío de Spark o usar Cuadernos de Amazon EMR para trabajar con Delta Lake en Amazon EMR.

spark-shell

1. Conéctese al nodo principal mediante SSH. Para obtener más información, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR.

2. Introduzca el siguiente comando para iniciar el shell de Spark. Para usar la PySpark carcasa, sustitúyala por`spark-shell`. `pyspark`

```
spark-shell \
   --conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension" \
   --conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

   Si utiliza Amazon EMR versión 6.15.0 o posteriores, también debe usar las siguientes configuraciones para usar un control de acceso detallado basado en Lake Formation con Delta Lake.

```
spark-shell \
  --conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
 \
  --conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

  --conf spark.sql.catalog.spark_catalog.lf.managed=true
```

spark-submit

1. Conéctese al nodo principal mediante SSH. Para obtener más información, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR.
2. Introduzca el siguiente comando para iniciar la sesión de Spark para Delta Lake.

```
spark-submit
–conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension"
–conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

   Si utiliza Amazon EMR versión 6.15.0 o posteriores, también debe usar las siguientes configuraciones para usar un control de acceso detallado basado en Lake Formation con Delta Lake.

```
spark-submit \  `
--conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
```

```
--conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

--conf spark.sql.catalog.spark_catalog.lf.managed=true
```

EMR Studio notebooks

Para iniciar una sesión de Spark con Amazon EMR Studio configure su sesión de Spark con el comando mágico %%configure de su cuaderno de Amazon EMR, como en el siguiente ejemplo. Para obtener más información, consulte Usar la magia de Cuadernos de Amazon EMR en la Guía de administración de Amazon EMR.

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions":  "io.delta.sql.DeltaSparkSessionExtension",
     "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog"
  }
}
```

Si utiliza Amazon EMR versión 6.15.0 o posteriores, también debe usar las siguientes configuraciones para usar un control de acceso detallado basado en Lake Formation con Delta Lake.

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions":
 "io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserver.connector.spark.sql
    "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog",
    "spark.sql.catalog.spark_catalog.lf.managed": "true"
  }
}
```

Escritura en una tabla de Delta Lake

El siguiente ejemplo muestra cómo crear un conjunto de datos de Delta Lake DataFrame y cómo escribirlo como tal. El ejemplo muestra cómo trabajar con conjuntos de datos con el intérprete

de comandos de Spark mientras está conectado al nodo principal usando SSH como usuario predeterminado de Hadoop.

> **Note**
>
> Para pegar muestras de código en el intérprete de comandos de Spark, escriba :paste en el símbolo del sistema, pegue el ejemplo y, a continuación, pulse CTRL + D.

PySpark

Spark incluye un intérprete de comandos basado en Python, `pyspark`, que puede utilizar para crear prototipos de programas de Spark escritos en Python. Al igual que con `spark-shell`, invoque `pyspark` en el nodo principal.

```
## Create a DataFrame
data =  spark.createDataFrame([("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")],
["id", "creation_date",  "last_update_time"])

## Write a DataFrame as a Delta Lake dataset to the S3  location
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string, creation_date
 string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");

data.writeTo("delta_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame
val data = Seq(("100",  "2015-01-01",  "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")).toDF("id", "creation_date",
 "last_update_time")
```

```
// Write a DataFrame as a Delta Lake dataset to the S3  location
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string,
creation_date string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");

data.write.format("delta").mode("append").saveAsTable("delta_table")
```

SQL

```
-- Create a Delta  Lake table with the S3 location
CREATE TABLE delta_table(id string,
creation_date string,
last_update_time string)
USING delta LOCATION
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table';

-- insert data into the table
INSERT INTO delta_table VALUES  ("100", "2015-01-01",
 "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01", "2015-01-01T13:51:40.519832Z");
```

Lectura de una tabla Delta Lake

PySpark

```
ddf = spark.table("delta_table")
ddf.show()
```

Scala

```
val ddf =  spark.table("delta_table")
ddf.show()
```

SQL

```
SELECT * FROM delta_table;
```

Uso de un clúster de Delta Lake con Spark y AWS Glue

Para usar el AWS Glue Catalog como almacén de tablas de Delta Lake, cree un clúster con los siguientes pasos. Para obtener información sobre cómo especificar la clasificación de Delta Lake mediante AWS Command Line Interface, consulte Proporcionar una configuración mediante la AWS Command Line Interface al crear un clúster o Proporcionar una configuración mediante el SDK de Java al crear un clúster.

Creación de un clúster de Delta Lake

1. Cree un archivo, `configurations.json`, con el siguiente contenido:

```
[{"Classification":"delta-defaults",
"Properties":{"delta.enabled":"true"}},
{"Classification":"spark-hive-site",
"Properties":
{"hive.metastore.client.factory.class":"com.amazonaws.glue.catalog.metastore.AWSGlueDataCat
```

2. Cree un clúster con la siguiente configuración y sustituya **`example Amazon S3 bucket path`** y **`subnet ID`** por sus propios valores.

```
aws emr create-cluster
    --release-label  emr-6.9.0
    --applications Name=Spark
    --configurations file://delta_configurations.json
    --region us-east-1
    --name My_Spark_Delta_Cluster
    --log-uri  s3://amzn-s3-demo-bucket/
    --instance-type m5.xlarge
    --instance-count 2
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Consideraciones y limitaciones

- Se admite el uso de Delta Lake con las versiones 6.9.0 y posteriores de Amazon EMR. Puede usar Apache Spark 3.x en clústeres de Amazon EMR con tablas Delta.

- Le recomendamos que utilice el esquema de URI s3 para las rutas de ubicación de S3 en lugar de s3a para obtener el mejor rendimiento, la mejor seguridad y la mejor fiabilidad. Para obtener más información, consulte Utilizar almacenamiento y sistemas de archivos.
- Con Amazon EMR 7.0, Delta Universal Format (UniForm) y no se convert-to-Iceberg admiten declaraciones.
- Con Amazon EMR 6.9 y 6.10, al almacenar los datos de la tabla de Delta Lake en Amazon S3, los datos de las columnas se convierten en `NULL` después de la operación de cambio de nombre de las columnas. Este problema se resuelve con Amazon EMR 6.11. Para obtener más información sobre la operación experimental de cambio de nombre de columnas, consulte la Column rename operation en la Guía del usuario de Delta Lake.
- Si crea una base de datos en el catálogo de datos de AWS Glue fuera de Apache Spark, la base de datos podría tener un `LOCATION` campo vacío. Como Spark no permite crear bases de datos con una propiedad de ubicación vacía, aparecerá el siguiente error si usa Spark en Amazon EMR para crear una tabla Delta en una base de datos de Glue y la base de datos tiene una propiedad `LOCATION` vacía:

  ```
  IllegalArgumentException: Can not create a Path from an empty string
  ```

  Para resolver este problema, cree la base de datos en el Catálogo de datos con una ruta válida y no vacía para el campo `LOCATION`. A fin de ver los pasos para implementar esta solución, consulte Excepción de argumento ilegal al crear una tabla en la Guía del usuario de Amazon Athena.

Historial de lanzamientos de Delta

En la siguiente tabla se muestra la versión de Delta incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Delta

Etiqueta de versión de Amazon EMR	Versión de Delta	Componentes instalados con Delta
emr-7.9.0	3.3.0-amzn-1	Not available.
emr-7.8.0	3.3.0-amzn-0	Not available.
emr-7.7.0	3.2.1-amzn-2	Not available.
emr-7.6.0	3.2.1-amzn-1	Not available.
emr-7.5.0	3.2.1-amzn-0	Not available.
emr-7.4.0	3.2.0-amzn-1	Not available.
emr-7.3.0	3.2.0-amzn-0	Not available.
emr-7.2.0	3.1.0	Not available.
emr-7.1.0	3.0.0	Not available.
emr-7.0.0	3.0.0	Not available.
emr-6.15.0	2.4.0	Not available.
emr-6.14.0	2.4.0	Not available.
emr-6.13.0	2.4.0	Not available.
emr-6.12.0	2.4.0	Not available.
emr-6.11.1	2.2.0	Not available.
emr-6.11.0	2.2.0	Not available.
emr-6.10.1	2.2.0	Not available.
emr-6.10.0	2.2.0	Not available.
emr-6.9.1	2.1.0	Not available.

Etiqueta de versión de Amazon EMR	Versión de Delta	Componentes instalados con Delta
emr-6.9.0	2.1.0	Not available.

Apache Flink

Apache Flink es un motor de flujo de datos de streaming que se puede usar para ejecutar el procesamiento de flujos en tiempo real con orígenes de datos de alto rendimiento. Flink admite la semántica temporal de out-of-order eventos para eventos, la semántica de una sola vez, el control de contrapresión y APIs está optimizado para escribir aplicaciones de streaming y por lotes.

Además, Flink dispone de conectores para orígenes de datos de terceros, tales como los siguientes:

- Amazon Kinesis Data Streams
- Apache Kafka
- Conector para Elasticsearch de Flink
- Twitter Streaming API
- Cassandra

Amazon EMR admite Flink como aplicación YARN para que pueda administrar los recursos junto con otras aplicaciones de un clúster. Flink-on-YARNIe permite enviar trabajos transitorios de Flink o puede crear un clúster de larga duración que acepte varios trabajos y asigne los recursos de acuerdo con la reserva general de YARN.

Flink se incluye en las versiones de Amazon EMR 5.1.0 y posteriores.

> **Note**
>
> La compatibilidad con la clase `FlinkKinesisConsumer` se agregó en la versión de lanzamiento de Amazon EMR 5.2.1.

En la siguiente tabla, se muestra la versión de Flink incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Flink.

Para ver la versión de los componentes instalados con Flink en esta versión, consulte Versiones de componentes de la versión 7.9.0.

Información sobre la versión de Flink para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-7.9.0	Flink 1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

En la tabla siguiente, se muestra la versión de Flink incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Flink.

Para ver la versión de los componentes instalados con Flink en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Flink para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.15.0	Flink 1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client,

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
		flink-jobmanager-config, hudi, delta-standalone-connectors

En la tabla siguiente, se muestra la versión de Flink incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Flink.

Para ver la versión de los componentes instalados con Flink en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Flink para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.36.2	Flink 1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Temas

- Creación de un clúster con Flink
- Configuración de Flink en Amazon EMR
- Trabajar con trabajos de Flink en Amazon EMR
- Uso del intérprete de comandos de Scala
- Encontrar la interfaz web de Flink
- Escalador automático de Flink

- Optimización de los tiempos de reinicio de los trabajos para las operaciones de escalado y recuperación de tareas
- Trabajar con trabajos de Flink de Zeppelin en Amazon EMR
- Historial de lanzamientos de Flink

Creación de un clúster con Flink

Puede lanzar un clúster con el AWS Management Console AWS CLI, o un AWS SDK.

Para lanzar un clúster con Flink instalado de la consola

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. En Software Configuration (Configuración de software), elija EMR Release emr-5.1.0 (EMR versión emr-5.1.0) o una versión posterior.
4. Elija Flink como aplicación, junto con las demás que desee instalar.
5. Seleccione las demás opciones que necesite y elija Create cluster (Crear clúster).

Para lanzar un clúster con Flink de la AWS CLI

- Cree el clúster con el siguiente comando:

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=MyKeyName,InstanceProfile=EMR_EC2_DefaultRole \
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args=flink-yarn-session,-d
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Configuración de Flink en Amazon EMR

Configuración de Flink con el metaalmacén de Hive y el Catálogo de Glue

Las versiones 6.9.0 y posteriores de Amazon EMR son compatibles con el metaalmacén de Hive y el Catálogo de AWS Glue con el conector de Apache Flink a Hive. En esta sección se describen los pasos necesarios para configurar el Catálogo de AWS Glue y el metaalmacén de Hive con Flink.

Temas

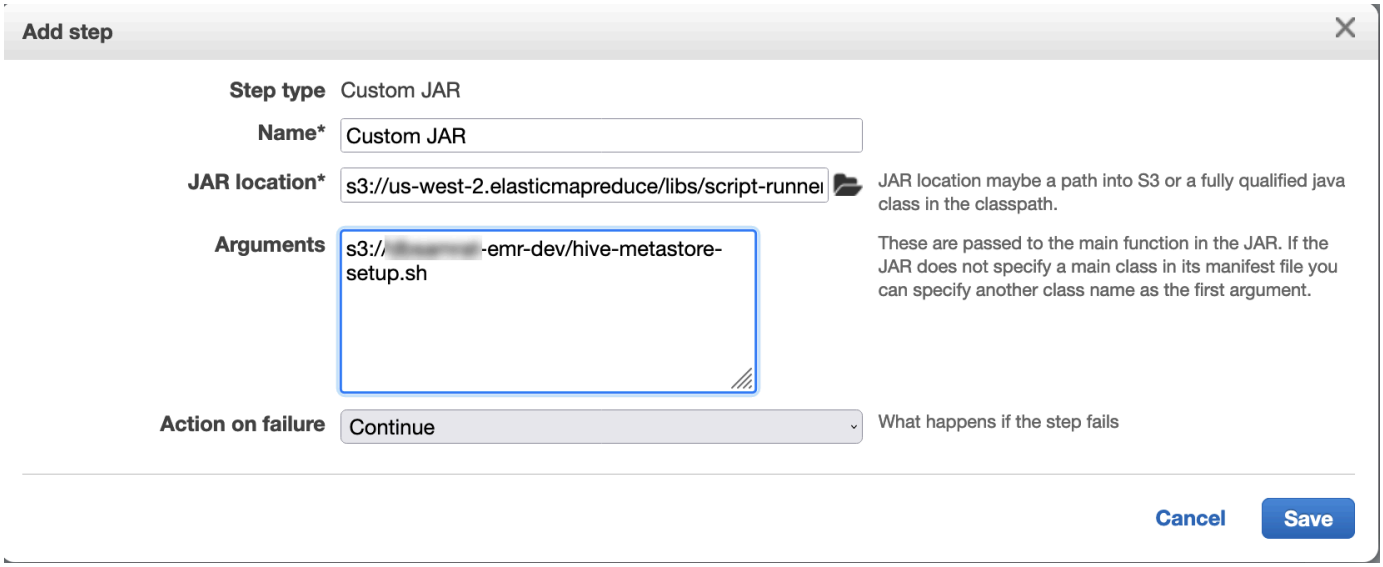
- Uso del metaalmacén de Hive
- Utilice el catálogo de datos de AWS Glue de Glue

Uso del metaalmacén de Hive

1. Cree un clúster de EMR con la versión 6.9.0 o posterior y al menos dos aplicaciones: Hive y Flink.
2. Utilice el ejecutor de scripts para ejecutar el siguiente script como una función escalonada:

   `hive-metastore-setup.sh`

```
sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
```

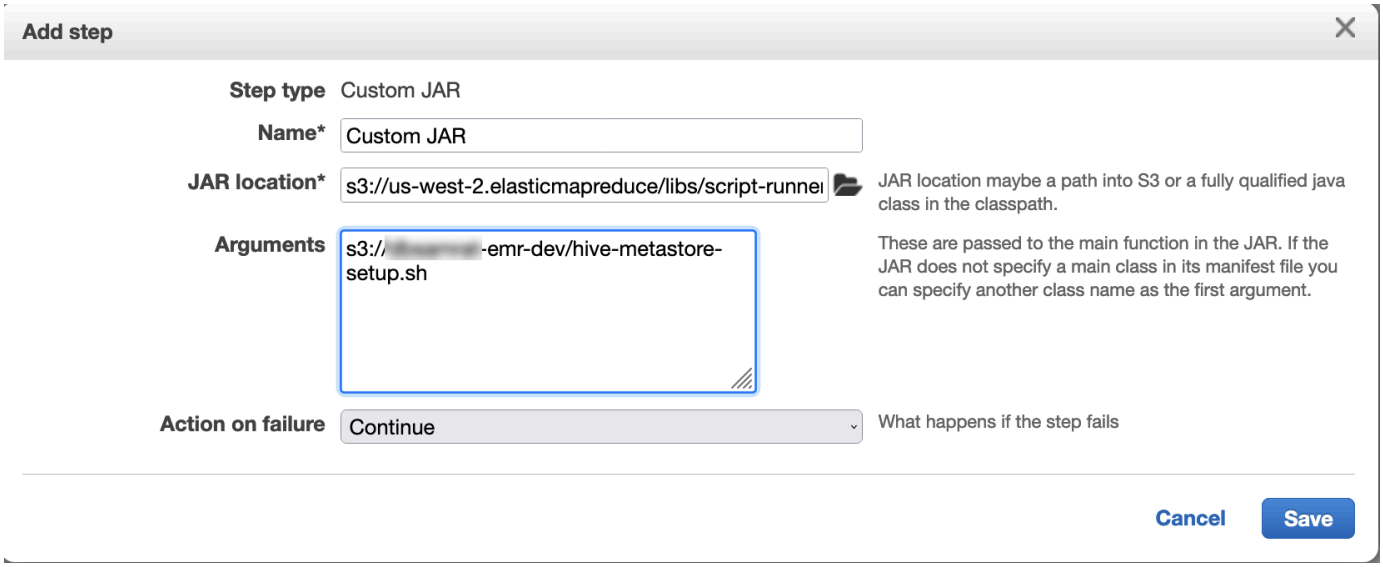


Utilice el catálogo de datos de AWS Glue de Glue

1. Cree un clúster de EMR con la versión 6.9.0 o posterior y al menos dos aplicaciones: Hive y Flink.
2. Seleccione Usar para metadatos de la tabla de Hive en la configuración del Catálogo de datos de AWS Glue para habilitar el Catálogo de datos en el clúster.
3. Utilice el ejecutor de scripts para ejecutar el siguiente script como función escalonada: Ejecución de comandos y scripts en un clúster de Amazon EMR:

   glue-catalog-setup.sh

   ```
   sudo cp /usr/lib/hive/auxlib/aws-glue-datacatalog-hive3-client.jar /usr/lib/flink/
   lib
   sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
   sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
   sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
   sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
   sudo chmod 755 /usr/lib/flink/lib/aws-glue-datacatalog-hive3-client.jar
   sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
   sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
   sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
   sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
   ```



Configuración de Flink con un archivo de configuración

Puede usar la API de configuración de Amazon EMR para configurar Flink con un archivo de configuración. Los archivos configurables dentro de la API son:

- `flink-conf.yaml`
- `log4j.properties`
- `flink-log4j-session`
- `log4j-cli.properties`

El archivo de configuración principal para Flink es `flink-conf.yaml`.

Para configurar el número de ranuras de tareas que se usan para Flink de la AWS CLI

1. Cree un archivo, `configurations.json`, con el siguiente contenido:

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "taskmanager.numberOfTaskSlots":"2"
      }
    }
```

```
]
```

2. A continuación, cree un clúster con la siguiente configuración:

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

> **Note**
>
> También puede cambiar algunas configuraciones con la API de Flink. Para obtener más información, consulte Conceptos en la documentación de Flink.
> Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Opciones de paralelismo

Como propietario de la aplicación, conoce bien los recursos que asignar a las tareas en Flink. Para los ejemplos de esta documentación, utilice el mismo número de tareas que el de las tareas de instancias que utilice para la aplicación. En general, recomendamos este nivel inicial de paralelismo, pero también puede aumentar el grado de detalle del paralelismo con ranuras de tareas, que generalmente no deben superar el número de núcleos virtuales por instancia. Para obtener más información sobre la arquitectura de Flink, consulte Conceptos en la documentación de Flink.

Configurar Flink en un clúster de EMR con varios nodos principales

El JobManager de Flink permanece disponible durante el proceso de conmutación por error del nodo principal en un clúster de Amazon EMR con varios nodos principales. A partir de Amazon

EMR 5.28.0, la JobManager alta disponibilidad también se activa automáticamente. No se necesita ninguna configuración manual.

Con las versiones 5.27.0 o anteriores de Amazon EMR, el error JobManager es único. Cuando se produce un JobManager error, pierde todos los estados de las tareas y no reanudará las tareas en ejecución. Puede habilitar la JobManager alta disponibilidad configurando el recuento de intentos de aplicación, los puntos de control y habilitándola ZooKeeper como almacenamiento de estado para Flink, como se muestra en el siguiente ejemplo:

```
[
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.resourcemanager.am.max-attempts": "10"
    }
  },
  {
    "Classification": "flink-conf",
    "Properties": {
        "yarn.application-attempts": "10",
        "high-availability": "zookeeper",
        "high-availability.zookeeper.quorum": "%{hiera('hadoop::zk')}",
        "high-availability.storageDir": "hdfs:///user/flink/recovery",
        "high-availability.zookeeper.path.root": "/flink"
    }
  }
]
```

Debe configurar tanto los intentos máximos maestros de aplicación para YARN como los intentos de aplicación para Flink. Para obtener más información, consulte Configuración de la alta disponibilidad de un clúster de YARN. También puede configurar los puntos de control de Flink para que los reiniciados JobManager recuperen los trabajos en ejecución a partir de puntos de control completados anteriormente. Para obtener más información, consulte Puntos de comprobación de Flink.

Configuración del tamaño del proceso de memoria

En el caso de las versiones de Amazon EMR que utilizan Flink 1.11.x, debe configurar el tamaño total del proceso de memoria para () y JobManager (`jobmanager.memory.process.size`) in. TaskManager `taskmanager.memory.process.size` `flink-conf.yaml` Para establecer

estos valores, puede configurar el clúster con la API de configuración o eliminar manualmente los comentarios de estos campos mediante SSH. Flink proporciona los siguientes valores predeterminados.

- `jobmanager.memory.process.size`: 1600 m
- `taskmanager.memory.process.size`: 1728 m

Para excluir el metaespacio y la sobrecarga de la JVM, utilice el tamaño total de la memoria de Flink (`taskmanager.memory.flink.size`) en lugar de `taskmanager.memory.process.size`. El valor predeterminado de `taskmanager.memory.process.size` es 1280 m. No se recomienda establecer `taskmanager.memory.process.size` y `taskmanager.memory.process.size`.

Todas las versiones de Amazon EMR que utilizan Flink 1.12.0 y versiones posteriores tienen los valores predeterminados enumerados en el conjunto de código abierto para Flink como valores predeterminados en Amazon EMR, por lo que no necesita configurarlos usted mismo.

Configuración del tamaño del archivo de salida del registro

Los contenedores de aplicaciones de Flink crean tres tipos de archivos de registro (archivos `.out`, archivos `.log` y archivos `.err`) y también escriben en ellos. Solo los archivos `.err` se comprimen y se eliminan del sistema de archivos, mientras que los archivos de registro `.log` y `.out` permanecen en el sistema de archivos. Para garantizar que estos archivos de salida se puedan administrar y que el clúster permanezca estable, puede configurar la rotación de registros en `log4j.properties` para establecer un número máximo de archivos y limitar sus tamaños.

Versiones 5.30.0 y posteriores de Amazon EMR

A partir de Amazon EMR 5.30.0, Flink utiliza el marco de registro log4j2 con el nombre de clasificación de configuración `flink-log4j.`. Con el siguiente ejemplo de configuración se muestra el formato log4j2.

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
      "appender.main.name": "MainAppender",
      "appender.main.type": "RollingFile",
      "appender.main.append" : "false",
      "appender.main.fileName" : "${sys:log.file}",
```

```
      "appender.main.filePattern" : "${sys:log.file}.%i",
      "appender.main.layout.type" : "PatternLayout",
      "appender.main.layout.pattern" : "%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p %-60c %x - %m
%n",
      "appender.main.policies.type" : "Policies",
      "appender.main.policies.size.type" : "SizeBasedTriggeringPolicy",
      "appender.main.policies.size.size" : "100MB",
      "appender.main.strategy.type" : "DefaultRolloverStrategy",
      "appender.main.strategy.max" : "10"
    },
  }
]
```

Versiones 5.29.0 y anteriores de Amazon EMR

Con las versiones 5.29.0 y anteriores de Amazon EMR, Flink usa el marco de registro log4j. El siguiente ejemplo ilustra la siguiente configuración de un dominio de log4j.

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
      "log4j.appender.file": "org.apache.log4j.RollingFileAppender",
      "log4j.appender.file.append":"true",
      # keep up to 4 files and each file size is limited to 100MB
      "log4j.appender.file.MaxFileSize":"100MB",
      "log4j.appender.file.MaxBackupIndex":4,
      "log4j.appender.file.layout":"org.apache.log4j.PatternLayout",
      "log4j.appender.file.layout.ConversionPattern":"%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p
 %-60c %x - %m%n"
    },
  }
]
```

Configuración de Flink para que se ejecute con Java 11

Las versiones 6.12.0 y posteriores de Amazon EMR proporcionan compatibilidad con el tiempo de ejecución de Java 11 para Flink. En las siguientes secciones, se describe cómo configurar el clúster para proporcionar compatibilidad con el tiempo de ejecución de Java 11 a Flink.

Temas

- Configuración de Flink para Java 11 al crear un clúster

- Configuración de Flink para Java 11 en un clúster en ejecución
- Confirmación del tiempo de ejecución de Java para Flink en un clúster en ejecución

Configuración de Flink para Java 11 al crear un clúster

Siga los pasos siguientes para crear un clúster de EMR con Flink y el tiempo de ejecución de Java 11. El archivo de configuración en el que se agrega la compatibilidad con el tiempo de ejecución de Java 11 es `flink-conf.yaml`.

Console

Para crear un clúster con Flink y el tiempo de ejecución de Java 11 en la consola

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Clústeres en EMR EC2 en el panel de navegación y, a continuación, Crear clúster.
3. Seleccione la versión 6.12.0 o posterior de Amazon EMR y elija instalar la aplicación de Flink. Seleccione cualquier otra aplicación que desee instalar en el clúster.
4. Continúe configurando el clúster. En la sección opcional Configuración de software, utilice la opción Ingresar la configuración predeterminada e ingrese la siguiente configuración:

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
]
```

5. Continúe con la configuración y el lanzamiento del clúster.

AWS CLI

Para crear un clúster con el tiempo de ejecución de Flink y Java 11 desde la CLI

1. Cree un archivo de configuración `configurations.json` para que Flink use Java 11.

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
]
```

2. Desde AWS CLI, cree un nuevo clúster de EMR con la versión 6.12.0 o posterior de Amazon EMR e instale la aplicación de Flink, como se muestra en el siguiente ejemplo:

```
aws emr create-cluster --release-label emr-6.12.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

Configuración de Flink para Java 11 en un clúster en ejecución

Siga los pasos siguientes para actualizar un clúster en ejecución de EMR con Flink y Java 11 en tiempo de ejecución. El archivo de configuración en el que se agrega la compatibilidad con el tiempo de ejecución de Java 11 es `flink-conf.yaml`.

Console

Para actualizar un clúster en ejecución con el tiempo de ejecución de Flink y Java 11 en la consola

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Clústeres en EMR activado EC2 en el panel de navegación y, a continuación, seleccione el clúster que desee actualizar.

> **Note**
>
> El clúster debe usar la versión 6.12.0 o posterior de Amazon EMR para admitir Java 11.

3. Seleccione la pestaña Configuraciones.
4. En la sección de Configuración del grupo de instancias, seleccione el grupo de instancias En ejecución que desea actualizar y, a continuación, seleccione Volver a configurar en el menú de acciones de la lista.
5. Vuelva a configurar el grupo de instancias con la opción Editar atributos de la siguiente manera. Seleccione Agregar nueva configuración después de cada uno de ellos.

Clasificación	Propiedad	Valor
`flink-conf`	`containerized.taskmanager.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`containerized.master.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`env.java.home`	`/usr/lib/jvm/jre-11`

6. Seleccione Guardar cambios para agregar los ajustes de configuración.

AWS CLI

Para actualizar un clúster en ejecución para usar el tiempo de ejecución de Flink y Java 11 en la CLI

Use el comando `modify-instance-groups` para especificar una nueva configuración para cada grupo de instancias en un clúster en ejecución.

1. Primero, cree un archivo de configuración `configurations.json` que configure a Flink para usar Java 11. En el siguiente ejemplo, *`ig-1xxxxxxx9`* reemplace por el ID del grupo de

instancias que desea volver a configurar. Guarde el siguiente archivo en el mismo directorio en el que ejecutará el comando `modify-instance-groups`.

```
[
   {
      "InstanceGroupId":"ig-1xxxxxxx9",
      "Configurations":[
         {
            "Classification":"flink-conf",
            "Properties":{
              "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "env.java.home":"/usr/lib/jvm/jre-11"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. Desde AWS CLI, ejecute el siguiente comando. Reemplace el ID del grupo de instancias que desea volver a configurar:

```
aws emr modify-instance-groups --cluster-id j-2AL4XXXXXX5T9 \
--instance-groups file://configurations.json
```

Confirmación del tiempo de ejecución de Java para Flink en un clúster en ejecución

A fin de determinar el tiempo de ejecución de Java para un clúster en ejecución, inicie sesión en el nodo principal con SSH, tal y como se describe en Conectarse al nodo principal utilizando SSH. A continuación, ejecute el siguiente comando:

```
ps -ef | grep flink
```

El comando `ps` con la opción `-ef` muestra una lista de todos los procesos en ejecución en el sistema. Puede filtrar esa salida con `grep` para encontrar menciones de la cadena `flink`. Revise el resultado para ver el valor del Entorno de ejecución de Java (JRE), `jre-XX`. En la siguiente salida, `jre-11` indica que se elige Java 11 en tiempo de ejecución para Flink.

```
flink    19130     1  0 09:17 ?        00:00:15 /usr/lib/jvm/jre-11/bin/
java -Djava.io.tmpdir=/mnt/tmp -Dlog.file=/usr/lib/flink/log/flink-flink-
historyserver-0-ip-172-31-32-127.log -Dlog4j.configuration=file:/usr/lib/flink/conf/
log4j.properties -Dlog4j.configurationFile=file:/usr/lib/flink/conf/log4j.properties
 -Dlogback.configurationFile=file:/usr/lib/flink/conf/logback.xml -classpath /usr/lib/
flink/lib/flink-cep-1.17.0.jar:/usr/lib/flink/lib/flink-connector-files-1.17.0.jar:/
usr/lib/flink/lib/flink-csv-1.17.0.jar:/usr/lib/flink/lib/flink-json-1.17.0.jar:/usr/
lib/flink/lib/flink-scala_2.12-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-java-
uber-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-scala-bridge_2.12-1.17.0.
```

Como alternativa, inicie sesión en el nodo principal con SSH e inicie una sesión de YARN de Flink con el comando `flink-yarn-session -d`. El resultado muestra la máquina virtual de Java (JVM) de Flink, `java-11-amazon-corretto` en el siguiente ejemplo:

```
2023-05-29 10:38:14,129 INFO  org.apache.flink.configuration.GlobalConfiguration
     [] - Loading configuration property: containerized.master.env.JAVA_HOME, /usr/lib/
jvm/java-11-amazon-corretto.x86_64
```

Trabajar con trabajos de Flink en Amazon EMR

Hay varias formas de interactuar con Flink en Amazon EMR: a través de la consola, la interfaz de Flink que se encuentra en la interfaz de usuario de seguimiento y en ResourceManager la línea de comandos. Puede enviar un archivo JAR a una aplicación de Flink con cualquiera de estas opciones. Una vez enviado un archivo JAR, se convierte en un trabajo gestionado por Flink. JobManager JobManager Se encuentra en el nodo YARN que aloja el daemon Application Master de la sesión de Flink.

Puede ejecutar una aplicación de Flink como trabajo de YARN en un clúster de ejecución prolongada o como clúster transitorio. En un clúster de ejecución prolongada, puede enviar varios trabajos de Flink a un clúster de Flink que se ejecuta en Amazon EMR. Si ejecuta un trabajo de Flink como clúster transitorio, su clúster de Amazon EMR solo existe durante el tiempo que se tarda en ejecutar la aplicación de Flink, por lo que solo se le cobrará por los recursos y el tiempo utilizado. Puede enviar un trabajo de Flink con la operación de la API Amazon `AddSteps` EMR, como argumento de paso de `RunJobFlow` la operación y mediante AWS CLI `add-steps` los `create-cluster` comandos o.

Inicio de una aplicación de YARN de Flink como un paso en un clúster de ejecución prolongada

Para iniciar una aplicación de Flink a la que varios clientes puedan enviar trabajos mediante las operaciones de la API de YARN, debe crear un clúster o agregar una aplicación de Flink a un clúster existente. Para obtener instrucciones sobre cómo crear un clúster nuevo, consulte Creación de un clúster con Flink. Para iniciar una sesión de YARN en un clúster existente, utilice los siguientes pasos desde la consola, la AWS CLI o el SDK de Java.

> **Note**
>
> El comando `flink-yarn-session` se agregó en Amazon EMR versión 5.5.0 como envoltorio para que el script `yarn-session.sh` simplifique la ejecución. Si utiliza una versión anterior de Amazon EMR, sustituya `bash -c "/usr/lib/flink/bin/yarn-session.sh -d"` por Argumentos en la consola o `Args` en el comando AWS CLI .

Para enviar un trabajo de Flink en un clúster existente desde la consola

Envíe la sesión de Flink de con el comando `flink-yarn-session` en un clúster existente.

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En la lista de clústeres, seleccione el clúster que lanzó con anterioridad.
3. En la página de detalles del clúster, elija Steps (Pasos), Add Step (Añadir paso).
4. Utilice las directrices que siguen para introducir los parámetros y, a continuación, elija Agregar.

Parámetro	Descripción
Step type (Tipo de paso)	JAR personalizada
Nombre	Un nombre que le ayuda a identificar el paso. Por ejemplo, *<example-flink-step-name>*.
Jar location (Ubicación de JAR)	command-runner.jar
Arguments	El comando `flink-yarn-session` con argumentos adecuados para su aplicació

Parámetro	Descripción
	n. Por ejemplo, *flink-yarn-session -d* inicia una sesión de Flink dentro de su clúster de YARN en un estado separado (). -d Consulte Configuración de YARN en la documentación de Flink más reciente para conocer los detalles de argumentos.

Para enviar un trabajo de Flink en un clúster existente con el AWS CLI

- Utilice el comando `add-steps` para agregar un trabajo de Flink a un clúster de ejecución prolongada. El siguiente comando de ejemplo especifica `Args="flink-yarn-session", "-d"` para iniciar una sesión de Flink dentro del clúster de YARN en un estado desconectado (`-d`). Consulte Configuración de YARN en la documentación de Flink más reciente para conocer los detalles de argumentos.

```
aws emr add-steps --cluster-id <j-XXXXXXXX> --steps Type=CUSTOM_JAR,Name=<example-flink-step-name>,Jar=command-runner.jar,Args="flink-yarn-session","-d"
```

Envío del trabajo a una aplicación de Flink existente en un clúster de ejecución prolongada

Si ya tiene una aplicación de Flink existente en un clúster de ejecución prolongada, puede especificar el ID de la aplicación de Flink del clúster para enviar el trabajo. Para obtener el ID de la aplicación, ejecute `yarn application -list` la operación de YarnClientAPI AWS CLI o a través de ella:

```
$ yarn application -list
16/09/07 19:32:13 INFO client.RMProxy: Connecting to ResourceManager at
 ip-10-181-83-19.ec2.internal/10.181.83.19:8032
Total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED,
 RUNNING]):1
Application-Id    Application-Name    Application-Type    User    Queue    State
 Final-State    Progress    Tracking-URL
application_1473169569237_0002    Flink session with 14 TaskManagers (detached)
   Apache Flink     hadoop    default            RUNNING          UNDEFINED
 100% http://ip-10-136-154-194.ec2.internal:33089
```

El ID de la aplicación para esta sesión de Flink es`application_1473169569237_0002`, que puedes usar para enviar trabajos a la aplicación desde el SDK AWS CLI o desde un SDK.

Example SDK para Java

```
List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("flink", "run", "-m", "yarn-cluster", "-yid",
 "application_1473169569237_0002", "-yn", "2", "/usr/lib/flink/examples/streaming/
WordCount.jar",
      "--input", "s3://amzn-s3-demo-bucket/pg11.txt", "--output", "s3://amzn-s3-demo-
bucket/alice2/");

StepConfig flinkRunWordCount = new StepConfig()
  .withName("Flink add a wordcount step")
  .withActionOnFailure("CONTINUE")
  .withHadoopJarStep(flinkWordCountConf);

stepConfigs.add(flinkRunWordCount);

AddJobFlowStepsResult res = emr.addJobFlowSteps(new AddJobFlowStepsRequest()
   .withJobFlowId("myClusterId")
   .withSteps(stepConfigs));
```

Example AWS CLI

```
aws emr add-steps --cluster-id <j-XXXXXXXX> \
--steps Type=CUSTOM_JAR,Name=Flink_Submit_To_Long_Running,Jar=command-runner.jar,\
Args="flink","run","-m","yarn-cluster","-yid","application_1473169569237_0002",\
"/usr/lib/flink/examples/streaming/WordCount.jar",\
"--input","s3://amzn-s3-demo-bucket/pg11.txt","--output","s3://amzn-s3-demo-bucket/
alice2/" \
--region <region-code>
```

Enviar un trabajo de Flink transitorio

El segundo ejemplo lanza un clúster transitorio que ejecuta un trabajo de Flink y, a continuación, termina al completarse.

Example SDK para Java

```
import java.util.ArrayList;
import java.util.List;
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;

public class Main_test {

 public static void main(String[] args) {
  AWSCredentials credentials_profile = null;
  try {
   credentials_profile = new ProfileCredentialsProvider("default").getCredentials();
  } catch (Exception e) {
   throw new AmazonClientException(
     "Cannot load credentials from .aws/credentials file. " +
       "Make sure that the credentials file exists and the profile name is specified
 within it.",
     e);
  }

  AmazonElasticMapReduce emr = AmazonElasticMapReduceClientBuilder.standard()
    .withCredentials(new AWSStaticCredentialsProvider(credentials_profile))
    .withRegion(Regions.US_WEST_1)
    .build();

  List<StepConfig> stepConfigs = new ArrayList<StepConfig>();
  HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("bash", "-c", "flink", "run", "-m", "yarn-cluster", "-yn", "2",
      "/usr/lib/flink/examples/streaming/WordCount.jar", "--input", "s3://path/to/
input-file.txt", "--output",
      "s3://path/to/output/");

  StepConfig flinkRunWordCountStep = new StepConfig()
    .withName("Flink add a wordcount step and terminate")
    .withActionOnFailure("CONTINUE")
    .withHadoopJarStep(flinkWordCountConf);
```

```
  stepConfigs.add(flinkRunWordCountStep);

  Application flink = new Application().withName("Flink");

  RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("flink-transient")
    .withReleaseLabel("emr-5.20.0")
    .withApplications(flink)
    .withServiceRole("EMR_DefaultRole")
    .withJobFlowRole("EMR_EC2_DefaultRole")
    .withLogUri("s3://path/to/my/logfiles")
    .withInstances(new JobFlowInstancesConfig()
      .withEc2KeyName("myEc2Key")
      .withEc2SubnetId("subnet-12ab3c45")
      .withInstanceCount(3)
      .withKeepJobFlowAliveWhenNoSteps(false)
      .withMasterInstanceType("m4.large")
      .withSlaveInstanceType("m4.large"))
    .withSteps(stepConfigs);

  RunJobFlowResult result = emr.runJobFlow(request);
  System.out.println("The cluster ID is " + result.toString());

 }

}
```

Example AWS CLI

Utilice el subcomando `create-cluster` para crear un clúster transitorio que termina cuando se completa el trabajo de Flink:

```
aws emr create-cluster --release-label emr-5.2.1 \
--name "Flink_Transient" \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--auto-terminate
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=<YourKeyName>,InstanceProfile=EMR_EC2_DefaultRole \
```

```
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args="bash","-c","\"flink run -m yarn-cluster /usr/lib/flink/examples/streaming/
WordCount.jar
--input s3://amzn-s3-demo-bucket/pg11.txt --output s3://amzn-s3-demo-bucket/alice/""
```

Uso del intérprete de comandos de Scala

El intérprete de comandos de Scala de Flink para clústeres de EMR solo está configurado para iniciar nuevas sesiones de YARN. Puede utilizar el shell Scala con el procedimiento siguiente.

Uso del intérprete de comandos de Scala de Flink en el nodo principal

1. Inicie sesión en el nodo principal con SSH, tal y como se describe en Conectarse al nodo principal utilizando SSH.
2. Escriba lo siguiente para iniciar un shell:

   En Amazon EMR versión 5.5.0 y versiones posteriores, puede usar el siguiente comando para iniciar un clúster de Yarn para el Scala Shell con uno. TaskManager

   ```
   % flink-scala-shell yarn 1
   ```

   En las versiones anteriores de Amazon EMR, utilice:

   ```
   % /usr/lib/flink/bin/start-scala-shell.sh yarn 1
   ```

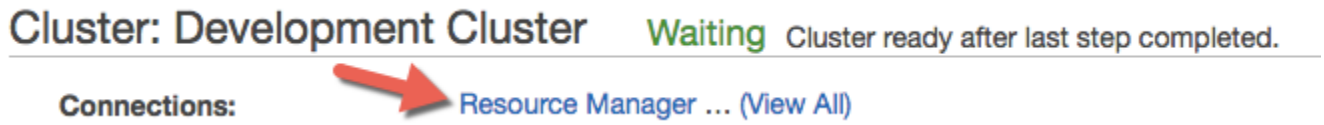
   Esto inicia el shell Scala de Flink para que pueda utilizar de forma interactiva Flink. De la misma forma que con otras interfaces y opciones, puede escalar el valor de opción `-n` utilizado en el ejemplo en función del número de tareas que desea ejecutar desde el shell.

   Para obtener más información, consulte Scala REPL en la documentación de Apache Flink.

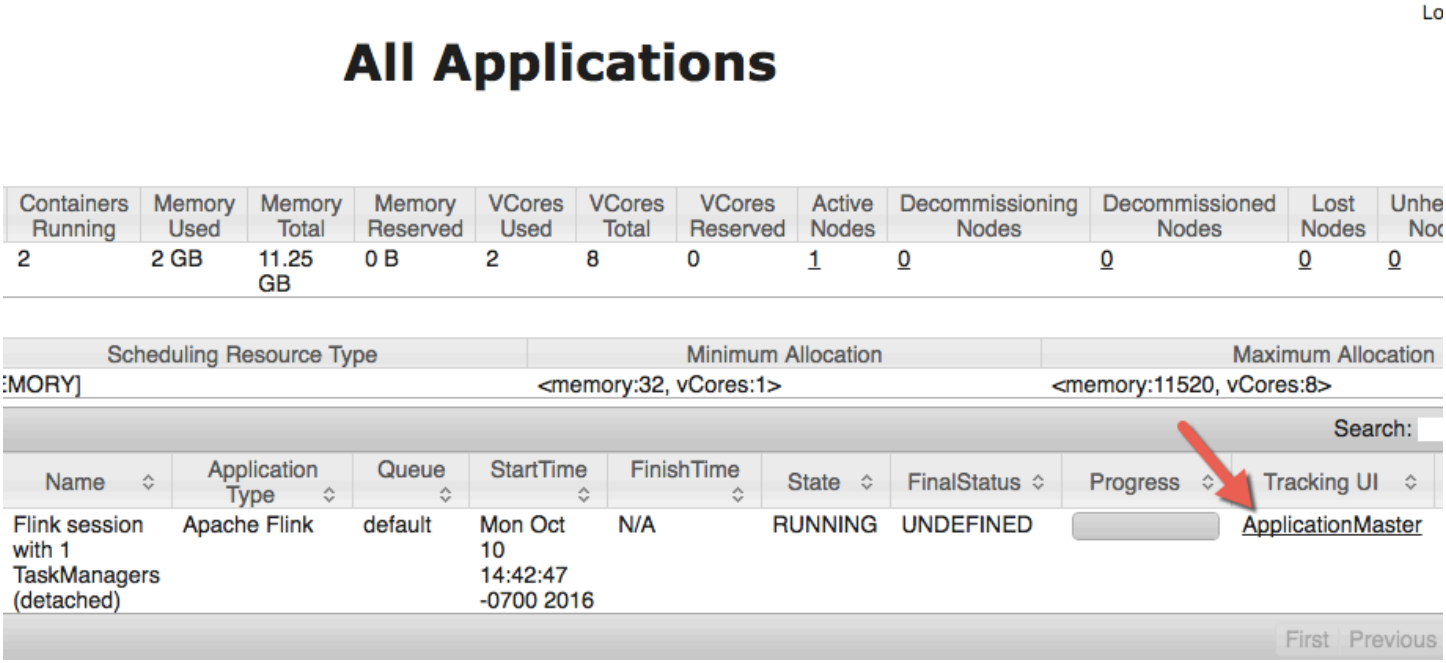
Encontrar la interfaz web de Flink

La aplicación maestra que pertenece a la aplicación de Flink aloja la interfaz web de Flink. Es una forma alternativa de enviar un JAR como trabajo o de ver el estado actual de otros trabajos. La interfaz web de Flink está activa siempre y cuando tenga una sesión de Flink en ejecución. Si ya tiene activo un trabajo de YARN de larga duración, puede seguir las instrucciones del tema Conectarse al nodo principal con SSH de la Guía de administración de Amazon EMR para

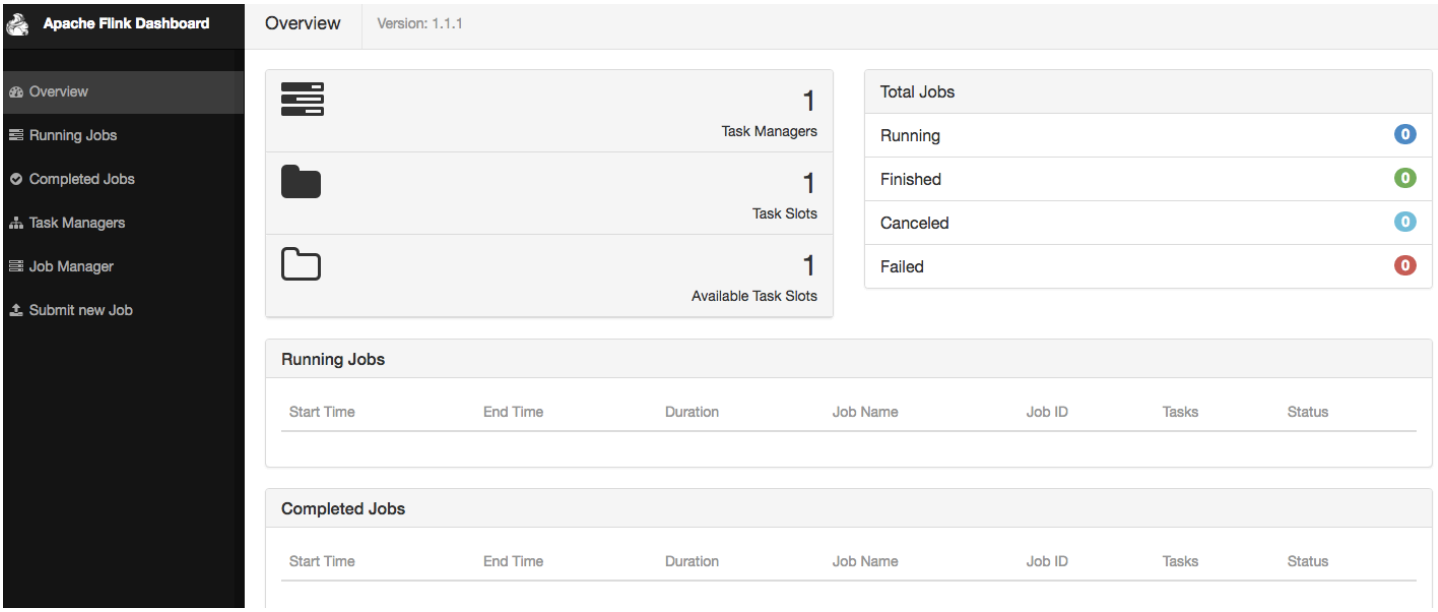
conectarse al YARN. ResourceManager Por ejemplo, si ha configurado un túnel SSH y ha activado un proxy en su navegador, debe elegir la ResourceManager conexión en Conexiones en la página de detalles del clúster de EMR.



Cuando encuentres el ResourceManager, selecciona la aplicación YARN que aloja una sesión de Flink. Elija el enlace bajo la columna Tracking UI (IU de seguimiento).



En la interfaz web de Flink, puede ver la configuración, enviar su propio JAR personalizado como trabajo o monitorizar los trabajos en curso.



Escalador automático de Flink

Descripción general

Las versiones 6.15.0 y posteriores de Amazon EMR admiten el escalador automático de Flink. La funcionalidad del escalador automático de trabajos recopila métricas de la ejecución de los trabajos de transmisión de Flink y escala automáticamente los vértices individuales de los trabajos. Esto reduce la contrapresión y cumple el objetivo de utilización establecido.

Para obtener más información, consulte la sección sobre el escalador automático de la documentación del operador de Kubernetes de Apache Flink.

Consideraciones

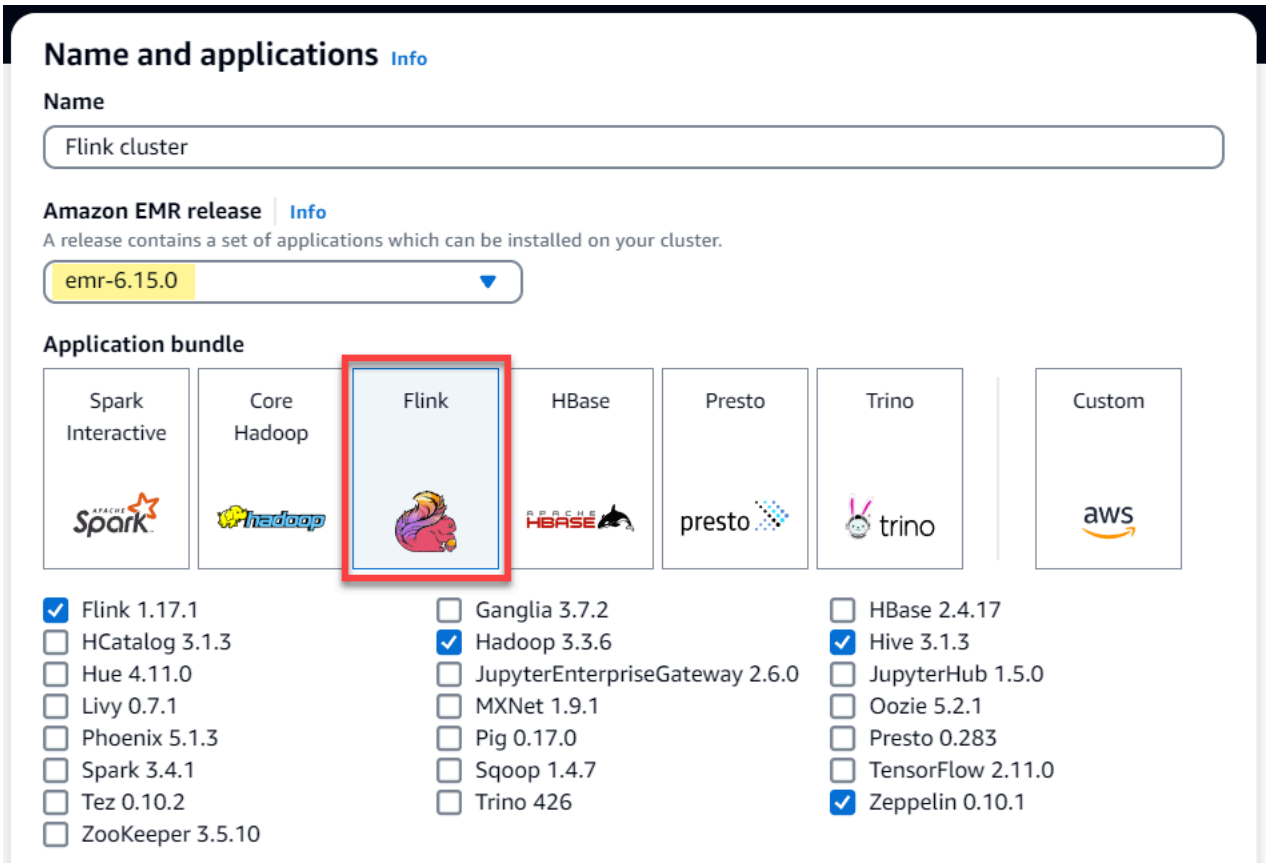
- Las versiones 6.15.0 y posteriores de Amazon EMR admiten el escalador automático de Flink.
- El escalador automático de Flink solo es compatible con trabajos de transmisión.
- Solo se admite el programador adaptativo. El programador predeterminado no es compatible.
- Se recomienda habilitar el escalado de clústeres para permitir el aprovisionamiento dinámico de recursos. Se prefiere el escalado administrado de Amazon EMR porque la evaluación de las métricas se realiza cada 5 a 10 segundos. En este intervalo, el clúster puede adaptarse más fácilmente al cambio en los recursos del clúster necesarios.

Habilitación del escalador automático

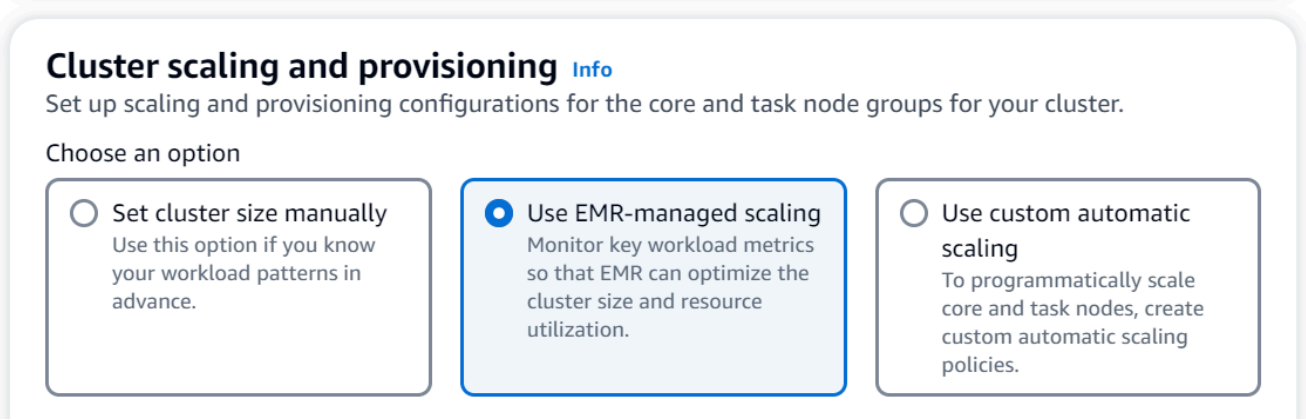
Siga los siguientes pasos para activar el escalador automático de Flink cuando cree un Amazon EMR en un clúster. EC2

1. En la consola de Amazon EMR, cree un nuevo clúster de EMR:

   a. Elija la versión `emr-6.15.0` de Amazon EMR o una posterior. Elija el paquete de aplicaciones de Flink y seleccione cualquier otra aplicación que desee incluir en su clúster.



   b. En la opción Escalado y aprovisionamiento del clúster, seleccione Usar escalado administrado de EMR.



2. En la sección Configuración de software, introduzca la siguiente configuración para habilitar el escalador automático de Flink. Para los escenarios de prueba, defina el intervalo de decisión, el intervalo de la ventana de métricas y el intervalo de estabilización en un valor inferior para que el trabajo tome inmediatamente una decisión de escalado a fin de facilitar la verificación.

```
[
  {
    "Classification": "flink-conf",
    "Properties": {
      "job.autoscaler.enabled": "true",
      "jobmanager.scheduler": "adaptive",
      "job.autoscaler.stabilization.interval": "60s",
      "job.autoscaler.metrics.window": "60s",
      "job.autoscaler.decision.interval": "10s",
      "job.autoscaler.debug.logs.interval": "60s"
    }
  }
]
```

3. Seleccione o configure cualquier otro ajuste como prefiera y cree el clúster con escalador automático de Flink.

Configuraciones del escalador automático

En esta sección, se describe la mayoría de las configuraciones que puede cambiar en función de sus necesidades específicas.

> **Note**
>
> En el caso de configuraciones que se basan en el tiempo como `time`, `interval` y `window`, la unidad predeterminada es milisegundos cuando no se especifica ninguna. Por lo tanto, un valor de `30` sin sufijo equivale a 30 milisegundos. Para otras unidades de tiempo, incluya el sufijo correspondiente: `s` para segundos, `m` para minutos y `h` para horas.

Temas

- Configuraciones de bucles del escalador automático
- Configuraciones de la agregación de métricas y el historial
- Configuraciones a nivel de vértice de trabajo
- Configuraciones de eventos de procesamiento pendientes
- Configuraciones de operaciones de escala

Configuraciones de bucles del escalador automático

El escalador automático busca las métricas a nivel de vértice de trabajo para cada intervalo de tiempo configurable, las convierte en variables procesables a escala, estima el nuevo paralelismo entre los vértices de trabajo y lo recomienda al programador de trabajos. Las métricas se recopilan solo después del tiempo de reinicio del trabajo y del intervalo de estabilización del clúster.

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
`job.autoscaler.enabled`	`false`	Habilite el escalado automático en su clúster de Flink.	`true`, `false`
`job.autoscaler.decision.interval`	`60s`	Intervalo de decisión del escalador automático.	`30` (la unidad predeterminada es milisegundos), `5m`, `1h`

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
`job.autoscaler.restart.time`	`3m`	Se utilizará el tiempo de reinicio previsto hasta que el operador pueda determinarlo de forma fiable a partir del historial.	`30` (la unidad predeterminada es milisegundos), `5m`, `1h`
`job.autoscaler.stabilization.interval`	`300s`	El periodo de estabilización en el que no se ejecutará ningún nuevo escalado.	`30` (la unidad predeterminada es milisegundos), `5m`, `1h`
`job.autoscaler.debug.logs.interval`	`300s`	El intervalo de registros de depuración del escalador automático.	`30` (la unidad predeterminada es milisegundos), `5m`, `1h`

Configuraciones de la agregación de métricas y el historial

El escalador automático busca las métricas, las agrega en una ventana deslizante basada en tiempo y las evalúa para tomar decisiones de escalado. El historial de decisiones de escalado de cada vértice de trabajo se utiliza para estimar el nuevo paralelismo. Estos valores tienen una fecha de vencimiento basada en la antigüedad y el tamaño del historial (al menos 1).

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
`job.autoscaler.metrics.window`	`600s`	`Scaling metrics aggregation window size.`	`30` (la unidad predeterminada es

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
			milisegundos), 5m, 1h
job.autoscaler.history.max.count	3	Número máximo de decisiones de escalado anteriores que se deben conservar por vértice.	De 1 a Integer.MAX_VALUE
job.autoscaler.history.max.age	24h	Número mínimo de decisiones de escalado anteriores que se deben conservar por vértice.	30 (la unidad predeterminada es milisegundos), 5m, 1h

Configuraciones a nivel de vértice de trabajo

El paralelismo de cada vértice de trabajo se modifica en función del objetivo de utilización y está limitado por los límites de paralelismo mínimo-máximo. No se recomienda fijar el objetivo de utilización cerca del 100 % (es decir, un valor de 1), y el límite de utilización funciona como amortiguador para gestionar las fluctuaciones de carga intermedias.

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
job.autoscaler.target.utilization	0.7	Objetivo de utilización del vértice.	0 - 1
job.autoscaler.target.utilization.boundary	0.4	Límite del objetivo de utilización del vértice. El escalado no se realizará si la velocidad de procesamiento actual está dentro de [target_rate /	0 - 1

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
		`(target_utilization - boundary)y(target_rate / (target_utilization + boundary)]` .	
`job.autoscaler.vertex.min-parallelism`	`1`	El paralelismo mínimo que puede utilizar el escalador automático.	`0` - `200`
`job.autoscaler.vertex.max-parallelism`	`200`	El paralelismo máximo que puede utilizar el escalador automático. Tenga en cuenta que se ignorará este límite si es superior al paralelismo máximo establecido en la configuración de Flink o directamente en cada operador.	`0` - `200`

Configuraciones de eventos de procesamiento pendientes

El vértice de trabajo necesita recursos adicionales para gestionar los eventos pendientes, o atrasados, que se acumulan durante el periodo de operación de escalado. Esto también se conoce como duración de `catch-up`. Si el tiempo de procesamiento del atraso supera el valor del `lag -threshold` configurado, el objetivo de utilización del vértice de trabajo aumenta hasta el nivel máximo. Esto ayuda a evitar operaciones de escalado innecesarias mientras se procesa la acumulación de eventos pendientes.

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
`job.autoscaler.backlog-processing.lag-threshold`	`5m`	Umbral de retraso que evitará escalados innecesarios y eliminará los mensajes pendientes responsables del retraso.	`30` (la unidad predeterminada es milisegundos), `5m`, `1h`
`job.autoscaler.catch-up.duration`	`15m`	El tiempo objetivo para procesar por completo cualquier acumulación después de una operación de escalado. Configúrelo en 0 para deshabilitar el escalado que se basa en la acumulación de eventos.	`30` (la unidad predeterminada es milisegundos), `5m`, `1h`

Configuraciones de operaciones de escala

El escalador automático no realiza la operación de reducción vertical inmediatamente después de una operación de escalado vertical dentro del periodo de gracia. Esto evita un ciclo innecesario de up-down-up-down operaciones de escalado provocado por fluctuaciones temporales de carga.

Podemos utilizar la relación de operaciones de reducción vertical para reducir gradualmente el paralelismo y liberar recursos a fin de hacer frente a los picos de carga temporales. Esto también ayuda a evitar una operación de escalado vertical menor innecesaria tras una operación de reducción vertical mayor.

Se puede detectar una operación de escalado vertical ineficaz según el historial de decisiones de escalado de vértices de trabajo anteriores para evitar nuevos cambios en el paralelismo.

Clave de configuración	Valor predeterminado	Descripción	Valores de ejemplo
`job.autoscaler.scale-up.grace-period`	`1h`	Duración durante la cual no se permite reducir verticalmente el tamaño de un vértice después de haberlo escalado verticalmente.	`30` (la unidad predeterminada es milisegundos), `5m`, `1h`
`job.autoscaler.scale-down.max-factor`	`0.6`	Factor máximo de reducción vertical. Un valor de `1` significa que no hay límite en la reducción vertical; `0.6` significa que el trabajo solo se puede reducir verticalmente con el 60 % del paralelismo original.	`0` - `1`
`job.autoscaler.scale-up.max-factor`	`100000.`	Relación máxima de escalado vertical. Un valor de `2.0` significa que el trabajo solo se puede escalar verticalmente con el 200 % del paralelismo actual.	`0` - `Integer.MAX_VALUE`
`job.autoscaler.scaling.effectiveness.detection.enabled`	`false`	Si se debe permitir la detección de operaciones de escalado ineficaces y permitir que el escalador automático bloquee nuevos escalados verticales.	`true`, `false`

Optimización de los tiempos de reinicio de los trabajos para las operaciones de escalado y recuperación de tareas

Cuando se produce un error en una tarea o se produce una operación de escalado, Flink intenta volver a ejecutar la tarea desde el último punto de control completado. El proceso de reinicio puede tardar un minuto o más en ejecutarse, según el tamaño del estado del punto de control y el número de tareas paralelas. Durante el periodo de reinicio, es posible que se acumulen tareas pendientes para el trabajo. Sin embargo, hay algunas formas en las que Flink optimiza la velocidad de los gráficos de recuperación y reinicio de la ejecución para mejorar la estabilidad del trabajo.

En esta página, se describen algunas de las formas en que Amazon EMR Flink puede mejorar el tiempo de reinicio de un trabajo durante las operaciones de recuperación o escalado de tareas.

Temas

- Recuperación local de tareas
- Punto de control incremental genérico basado en registros
- Recuperación detallada
- Mecanismo de reinicio combinado en el programador adaptativo

Recuperación local de tareas

> **Note**
>
> La recuperación local de tareas es compatible a partir de la versión 6.0.0 de Amazon EMR.

Con los puntos de control de Flink, cada tarea produce una instantánea de su estado que Flink graba en un almacenamiento distribuido como Amazon S3. En los casos de recuperación, las tareas restauran su estado desde el almacenamiento distribuido. El almacenamiento distribuido ofrece tolerancia a errores y puede redistribuir el estado durante el reescalado, ya que todos los nodos pueden acceder a él.

Sin embargo, un almacén distribuido remoto también tiene una desventaja: todas las tareas deben leer su estado desde una ubicación remota a través de la red. Esto puede provocar tiempos de recuperación prolongados para los estados de gran tamaño durante las operaciones de recuperación o escalado de tareas.

Este problema del tiempo de recuperación prolongado se resuelve mediante la recuperación local de tareas. Las tareas escriben su estado en el punto de control de un almacenamiento secundario que es local para la tarea, por ejemplo, en un disco local. También almacenan su estado en el almacenamiento principal, o Amazon S3 en nuestro caso. Durante la recuperación, el programador programa las tareas en el mismo administrador de tareas en el que se ejecutaron anteriormente, de modo que se puedan recuperar del almacén de estado local en lugar de leerlas del almacén de estado remoto. Para obtener más información, consulte Task-Local Recovery en la documentación de Apache Flink.

Nuestras pruebas comparativas con trabajos de muestra han demostrado que el tiempo de recuperación se ha reducido de unos minutos a unos pocos segundos con la opción de recuperación local de tareas habilitada.

Para habilitar la recuperación local de tareas, defina las siguientes configuraciones en el archivo `flink-conf.yaml`. Especifique el valor del intervalo de puntos de control en milisegundos.

```
state.backend.local-recovery: true
state.backend: hasmap or rocksdb
state.checkpoints.dir: s3://storage-location-bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

Punto de control incremental genérico basado en registros

> **Note**
>
> Los puntos de control incrementales genéricos basados en registros son compatibles a partir de la versión 6.10.0 de Amazon EMR.

En Flink 1.16, se agregaron puntos de control incrementales genéricos basados en registros para mejorar la velocidad de los puntos de control. Un intervalo de puntos de control más rápido suele reducir el trabajo de recuperación, ya que es necesario volver a procesar menos eventos después de la recuperación. Para obtener más información, consulte Improving speed and stability of checkpointing with generic log-based incremental checkpoints en el blog de Apache Flink.

Con trabajos de muestra, nuestras pruebas comparativas han demostrado que el tiempo de los puntos de control se ha reducido de unos minutos a unos pocos segundos con el punto de control incremental genérico basado en registros.

Para habilitar los puntos de control incrementales genéricos basados en registros, defina las siguientes configuraciones en el archivo `flink-conf.yaml`. Especifique el valor del intervalo de puntos de control en milisegundos.

```
    state.backend.changelog.enabled: true
    state.backend.changelog.storage: filesystem
    dstl.dfs.base-path: s3://bucket-path/changelog
    state.backend.local-recovery: true
    state.backend: rocksdb
    state.checkpoints.dir: s3://bucket-path/checkpoint
    execution.checkpointing.interval: 15000
```

Recuperación detallada

> **Note**
>
> El soporte de la recuperación detallada para el programador predeterminado está disponible a partir de la versión 6.0.0 de Amazon EMR. El soporte de la recuperación detallada para el programador predeterminado está disponible a partir de la versión 6.15.0 de Amazon EMR.

Cuando se produce un error en una tarea durante la ejecución, Flink restablece todo el gráfico de ejecución y activa una nueva ejecución completa desde el último punto de control completado. Esto es más caro que volver a ejecutar las tareas con errores. La recuperación detallada reinicia solo el componente conectado a la canalización de la tarea con el error. En el siguiente ejemplo, el gráfico de tareas tiene 5 vértices (de `A` a `E`). Todas las conexiones entre los vértices se canalizan con una distribución puntual y el valor `parallelism.default` para el trabajo se establece en 2.

```
A # B # C # D # E
```

En este ejemplo, hay un total de 10 tareas en ejecución. La primera tubería (`a1haciae1`) pasa por un TaskManager (`TM1`), y la segunda canalización (`a2haciae2`) corre por otra TaskManager (`TM2`).

```
a1 # b1 # c1 # d1 # e1
a2 # b2 # c2 # d2 # e2
```

Hay dos componentes conectados por canalización: `a1 # e1` y `a2 # e2`. Si `TM1` alguna de las dos `TM2` falla, la falla solo afecta a las 5 tareas en proceso, donde TaskManager estaba ejecutándose. La estrategia de reinicio solo inicia el componente canalizado afectado.

La recuperación detallada solo funciona con trabajos de Flink perfectamente paralelos. No es compatible con las operaciones `keyBy()` o `redistribute()`. Para obtener más información, consulte FLIP-1: Fine Grained Recovery from Task Failures en el proyecto de Jira Flink Improvement Proposal.

Para habilitar la recuperación detallada, establezca las siguientes configuraciones en el archivo `flink-conf.yaml`.

```
jobmanager.execution.failover-strategy: region
restart-strategy: exponential-delay or fixed-delay
```

Mecanismo de reinicio combinado en el programador adaptativo

> **Note**
>
> El mecanismo de reinicio combinado del programador adaptativo es compatible a partir de la versión 6.15.0 de Amazon EMR.

El programador adaptativo puede ajustar el paralelismo del trabajo en función de los espacios disponibles. Reduce automáticamente el paralelismo si no hay suficientes espacios disponibles para ajustarse al paralelismo del trabajo configurado. Si hay nuevos espacios disponibles, la tarea se escala verticalmente de nuevo al paralelismo del trabajo configurado. Un programador adaptativo evita el tiempo de inactividad del trabajo cuando no hay suficientes recursos disponibles. Este es el programador compatible con el escalador automático de Flink. Recomendamos el programador adaptativo con Amazon EMR Flink por estos motivos. Sin embargo, los programadores adaptativos pueden realizar varios reinicios en un periodo corto, uno por cada nuevo recurso agregado. Esto puede provocar una disminución en el rendimiento del trabajo.

Con Amazon EMR 6.15.0 y versiones posteriores, Flink cuenta con un mecanismo de reinicio combinado en el programador adaptativo que abre una ventana de reinicio cuando se agrega el primer recurso y, a continuación, espera hasta alcanzar el intervalo de ventana configurado del minuto predeterminado. Realiza un único reinicio cuando hay suficientes recursos disponibles para ejecutar el trabajo con el paralelismo configurado o cuando se agota el intervalo.

Con trabajos de muestra, nuestras pruebas comparativas han demostrado que esta característica procesa un 10 % de los registros más que el comportamiento predeterminado cuando se utiliza el programador adaptativo y el escalador automático de Flink.

Para habilitar el mecanismo de reinicio combinado, defina las siguientes configuraciones en el archivo `flink-conf.yaml`.

```
jobmanager.adaptive-scheduler.combined-restart.enabled: true
jobmanager.adaptive-scheduler.combined-restart.window-interval: 1m
```

Trabajar con trabajos de Flink de Zeppelin en Amazon EMR

Introducción

Las versiones 6.10.0 y posteriores de Amazon EMR admiten la integración de Apache Zeppelin con Apache Flink. Puede enviar trabajos de Flink de forma interactiva a través de los cuadernos de Zeppelin. Con el intérprete de Flink, puede ejecutar consultas de Flink, definir los trabajos de streaming y lotes de Flink y visualizar el resultado en los cuadernos de Zeppelin. El intérprete de Flink se basa en la API de REST de Flink. Esto le permite acceder a los trabajos de Flink y manipularlos desde el entorno de Zeppelin para procesar y analizar los datos en tiempo real.

Hay cuatro subintérpretes en el intérprete de Flink. Tienen diferentes propósitos, pero todos están en la JVM y comparten los mismos puntos de entrada a Flink preconfigurados (`ExecutionEnviroment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment`). Los intérpretes son los siguientes:

- `%flink`: crea `ExecutionEnvironment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment` y proporciona un entorno de Scala.
- `%flink.pyflink`: proporciona un entorno de Python.
- `%flink.ssql`: proporciona un entorno de SQL de streaming.
- `%flink.bsql`: proporciona un entorno de SQL de lotes.

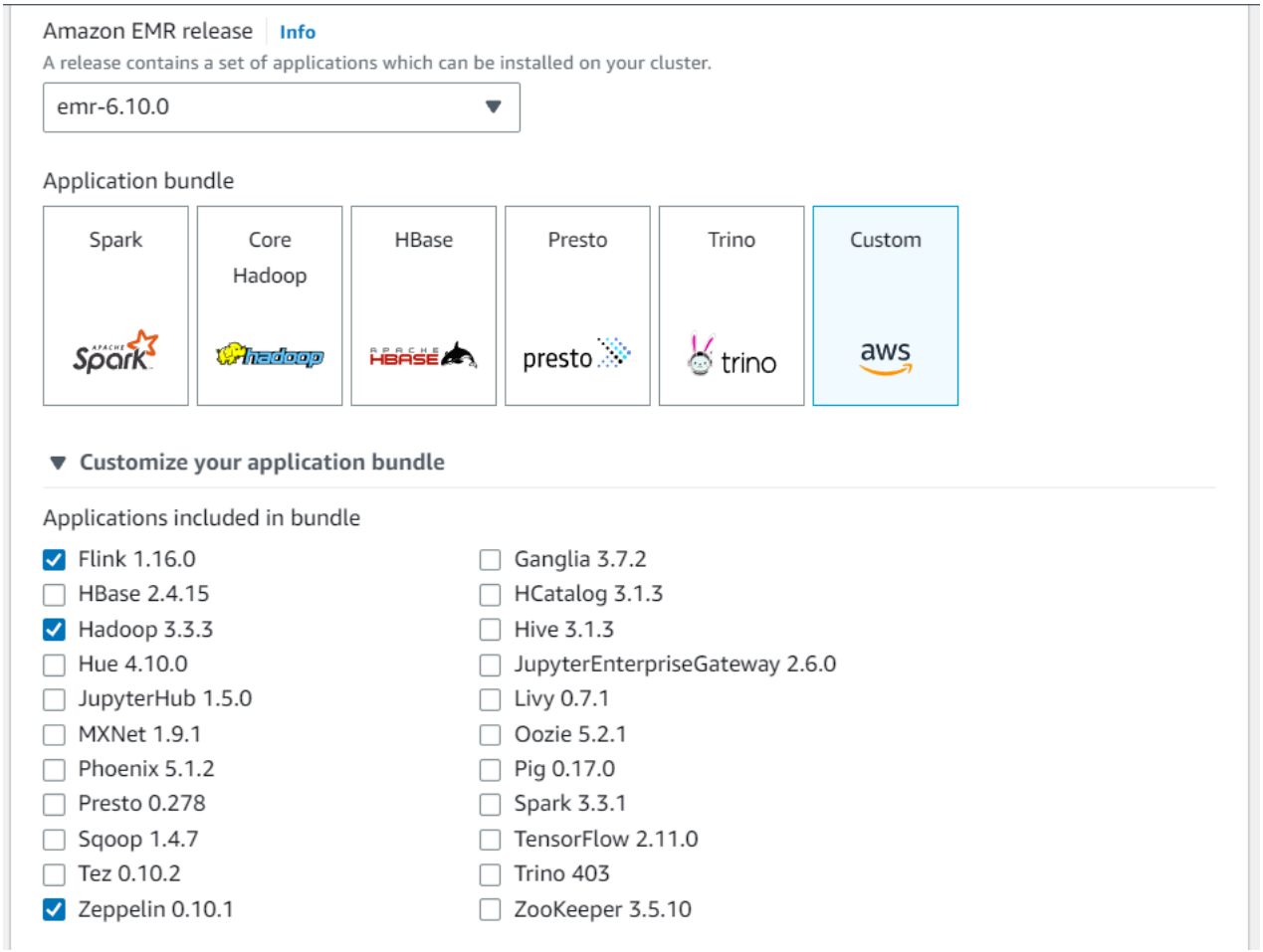
Requisitos previos

- La integración de Zeppelin con Flink es compatible con los clústeres creados con la versión 6.10.0 y posteriores de Amazon EMR.
- Para ver las interfaces web que están alojadas en los clústeres de EMR como se requiere en estos pasos, debe configurar un túnel de SSH para permitir el acceso entrante. Para obtener más información, consulte Configurar ajustes de proxy para ver sitios web alojados en el nodo principal.

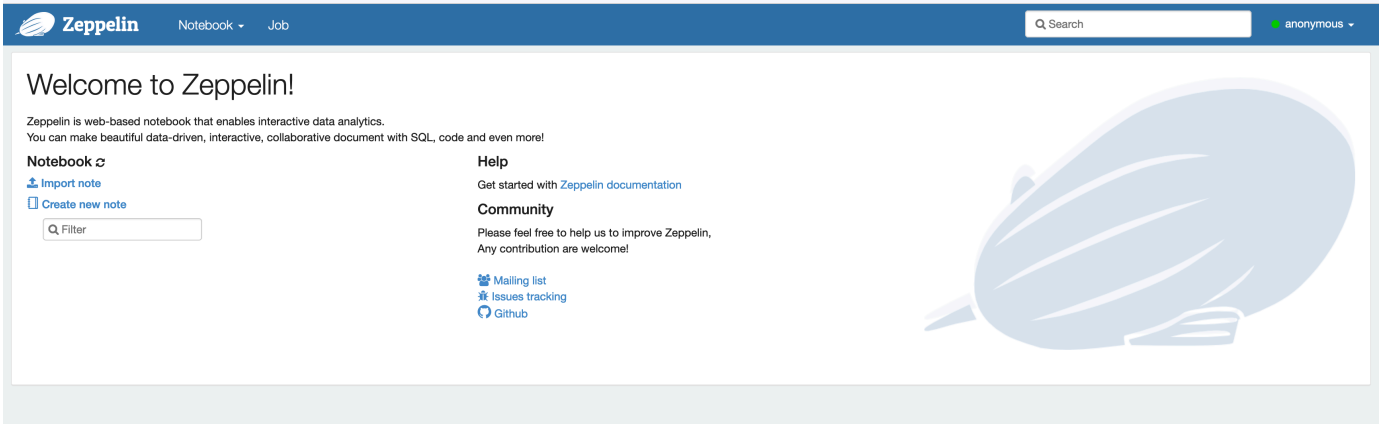
Configuración de Zeppelin-Flink en un clúster de EMR

Siga los siguientes pasos para configurar Apache Flink en Apache Zeppelin para que se ejecute en un clúster de EMR:

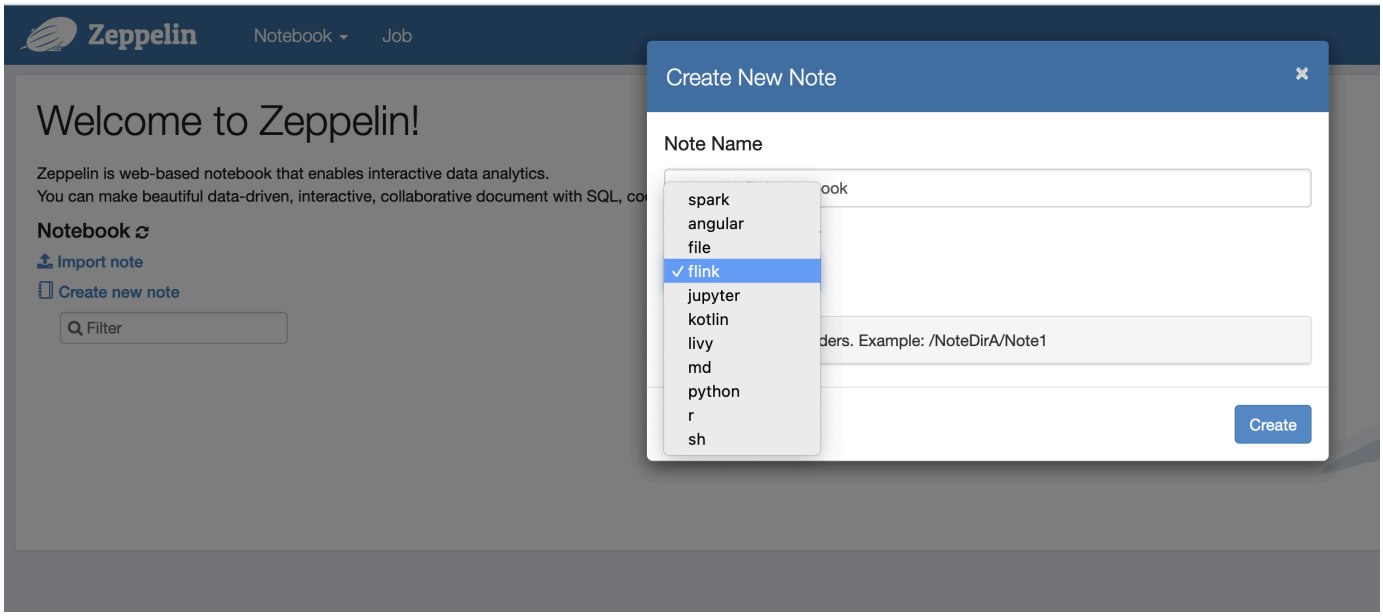
1. Cree un clúster nuevo desde la consola de Amazon EMR. Seleccione emr-6.10.0 o una versión posterior para la versión de Amazon EMR. A continuación, elija personalizar el paquete de aplicaciones con la opción Personalizado. Incluya al menos Flink, Hadoop y Zeppelin en su paquete.



2. Cree el resto del clúster con la configuración que prefiera.
3. Una vez que el clúster esté en ejecución, selecciónelo en la consola para ver sus detalles y abrir la pestaña Aplicaciones. Seleccione Zeppelin en la sección Interfaces de usuario de aplicaciones para abrir la interfaz web de Zeppelin. Asegúrese de haber configurado el acceso a la interfaz web de Zeppelin con un túnel de SSH al nodo principal y una conexión de proxy, tal y como se describe en Requisitos previos.



4. Ahora puede crear una nota nueva en un cuaderno de Zeppelin con Flink como intérprete predeterminado.



5. Consulte los siguientes ejemplos de código que muestran cómo ejecutar trabajos de Flink desde un cuaderno de Zeppelin.

Ejecución de trabajos de Flink con Zeppelin-Flink en un clúster de EMR

- Ejemplo 1: Flink Scala

  a) WordCount Ejemplo de lote (SCALA)

```
%flink

val data = benv.fromElements("hello world", "hello flink", "hello hadoop")
```

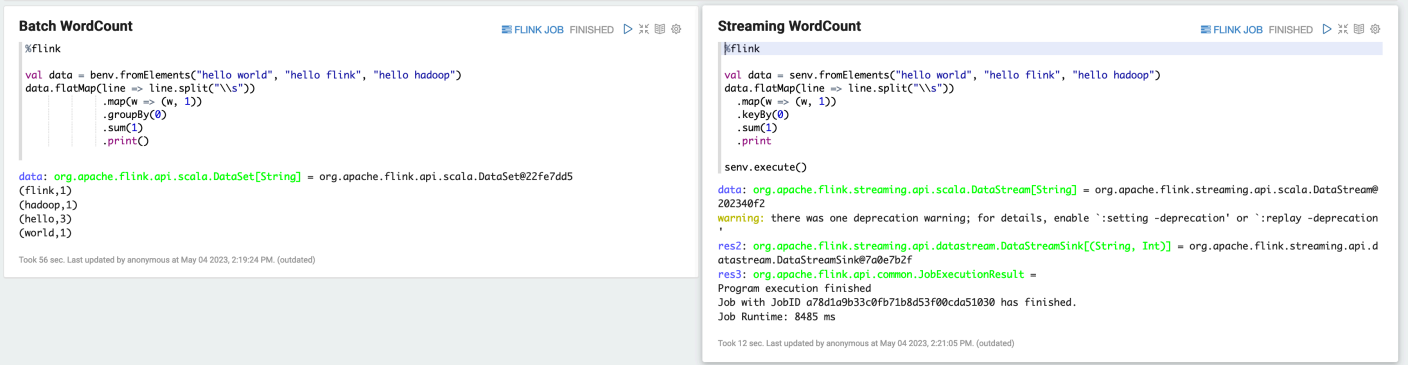
```
data.flatMap(line => line.split("\\s"))
             .map(w => (w, 1))
             .groupBy(0)
             .sum(1)
             .print()
```

b) WordCount Ejemplo de transmisión (SCALA)

```
%flink

val data = senv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
  .map(w => (w, 1))
  .keyBy(0)
  .sum(1)
  .print

senv.execute()
```



- Ejemplo 2: SQL en streaming de Flink

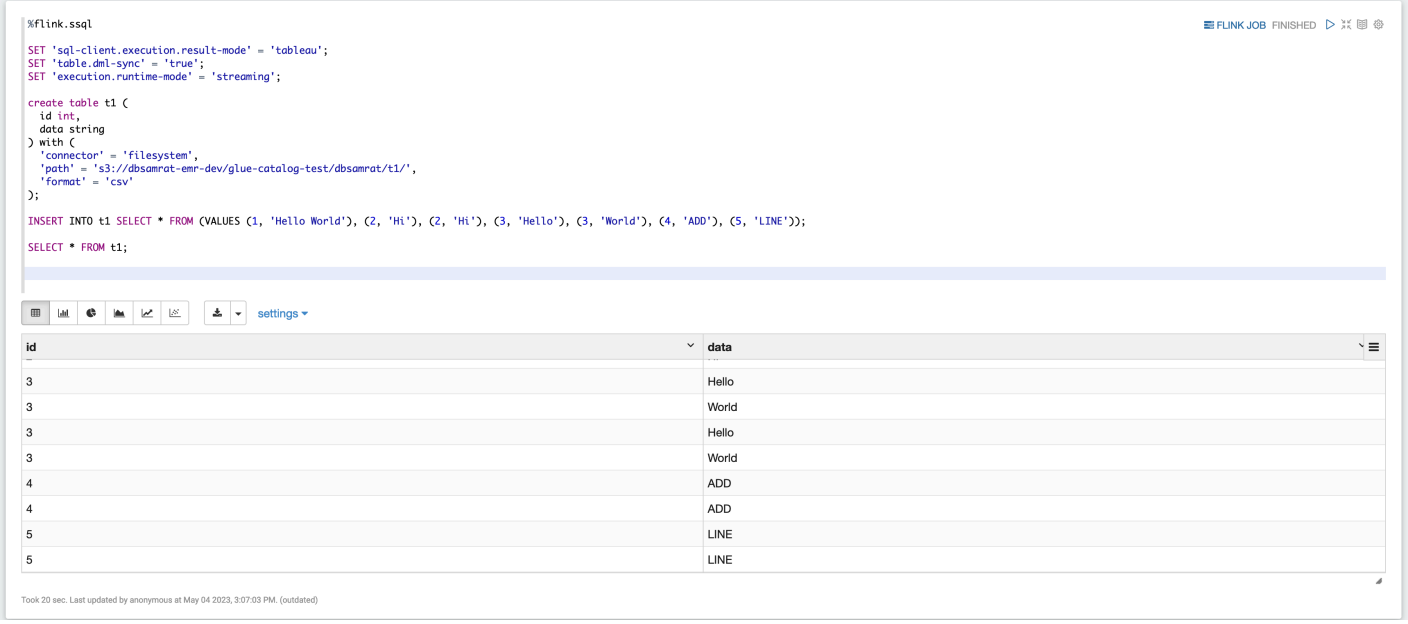
```
%flink.ssql
SET 'sql-client.execution.result-mode' = 'tableau';
SET 'table.dml-sync' = 'true';
SET 'execution.runtime-mode' = 'streaming';

create table dummy_table (
  id int,
  data string
) with (
  'connector' = 'filesystem',
  'path' = 's3://s3-bucket/dummy_table',
  'format' = 'csv'
```

```
);

INSERT INTO dummy_table SELECT * FROM (VALUES (1, 'Hello World'), (2, 'Hi'), (2,
 'Hi'), (3, 'Hello'), (3, 'World'), (4, 'ADD'), (5, 'LINE'));

SELECT * FROM dummy_table;
```



- Ejemplo 3: Pyflink. Tenga en cuenta que debe cargar su propio archivo de texto de muestra `word.txt` con un nombre en su bucket de S3.

```
%flink.pyflink

import argparse
import logging
import sys

from pyflink.common import Row
from pyflink.table import (EnvironmentSettings, TableEnvironment, TableDescriptor,
 Schema,
                           DataTypes, FormatDescriptor)
from pyflink.table.expressions import lit, col
from pyflink.table.udf import udtf

def word_count(input_path, output_path):
    t_env = TableEnvironment.create(EnvironmentSettings.in_streaming_mode())
    # write all the data to one file
    t_env.get_config().set("parallelism.default", "1")
```

```
    # define the source
    if input_path is not None:
        t_env.create_temporary_table(
            'source',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .build())
                           .option('path', input_path)
                           .format('csv')
                           .build())
        tab = t_env.from_path('source')
    else:
        print("Executing word_count example with default input data set.")
        print("Use --input to specify file input.")
        tab = t_env.from_elements(map(lambda i: (i,), word_count_data),
                                  DataTypes.ROW([DataTypes.FIELD('line',
DataTypes.STRING())]))

    # define the sink
    if output_path is not None:
        t_env.create_temporary_table(
            'sink',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .option('path', output_path)
                           .format(FormatDescriptor.for_format('canal-json')
                                   .build())
                           .build())
    else:
        print("Printing result to stdout. Use --output to specify output path.")
        t_env.create_temporary_table(
            'sink',
            TableDescriptor.for_connector('print')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .build())
```

```
    @udtf(result_types=[DataTypes.STRING()])
    def split(line: Row):
        for s in line[0].split():
            yield Row(s)

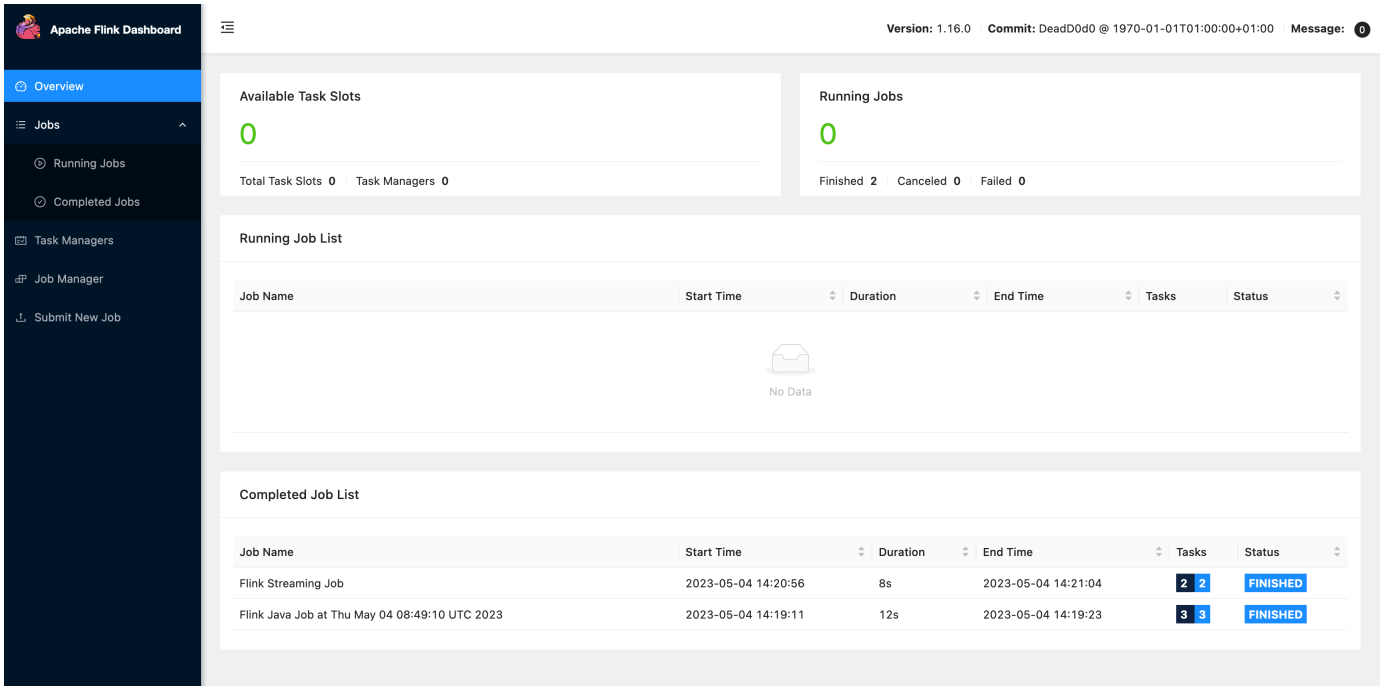
    # compute word count
    tab.flat_map(split).alias('word') \
       .group_by(col('word')) \
       .select(col('word'), lit(1).count) \
       .execute_insert('sink') \
       .wait()


logging.basicConfig(stream=sys.stdout, level=logging.INFO, format="%(message)s")


word_count("s3://s3_bucket/word.txt", "s3://s3_bucket/demo_output.txt")
```

1. Elija FLINK JOB en la interfaz de usuario de Zeppelin para acceder a la interfaz de usuario web de Flink y verla.

2. Al elegir FLINK JOB, se accede a la consola web de Flink en otra pestaña del navegador.



Historial de lanzamientos de Flink

En la siguiente tabla se muestra la versión de Flink incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Flink

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-7.9.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client,

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
		flink-jobmanager-config, hudi, delta
emr-7.8.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.7.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-7.6.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.5.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-7.4.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.3.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-7.2.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-5.36.2	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-7.1.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-7.0.0	1.18.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.15.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.14.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.13.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.12.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.11.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.11.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.10.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.10.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.9.1	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.9.0	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.8.1	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.8.0	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.7.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-5.36.1	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.36.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.6.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-5.35.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.5.0	1.14.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.4.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.3.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.3.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-6.1.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-6.1.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.34.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.33.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.33.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.32.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.32.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.30.2	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.1	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.0	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.29.0	1.9.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.1	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.0	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.27.1	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.27.0	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.26.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.25.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.1	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.23.1	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.23.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.22.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.21.2	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.1	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.0	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.20.1	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.20.0	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.19.1	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.19.0	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.1	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.0	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.17.2	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.1	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.0	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.16.1	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.16.0	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.15.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.15.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.2	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.14.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.12.3	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.2	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.12.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.4	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.3	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.11.2	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.10.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.10.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.9.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.9.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.3	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.2	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.8.1	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.0	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.7.1	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.7.0	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.1	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.0	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.5.4	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.3	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.2	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.5.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.4.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.4.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.2	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.1	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.3.0	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.3	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.2	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.2.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.1.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Etiqueta de versión de Amazon EMR	Versión de Flink	Componentes instalados con Flink
emr-5.1.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Notas de la versión de Flink por versión

Consulte las siguientes secciones para ver las notas completas de la versión.

Amazon EMR 7.9.0 - Notas de la versión de Flink

Amazon EMR 7.9.0: cambios en Flink

Tipo	Descripción
Nueva característica	A partir de Amazon EMR 7.9.0, Apache Flink out-of-the-box admite los formatos de archivo Avro, Parquet y ORC. Puede usar estos formatos directamente con cualquier API de Flink (DataStream, tabla o SQL) sin necesidad de ninguna configuración adicional.
Nueva característica	A partir de la versión 7.9.0 de Amazon EMR, puede habilitar los catálogos de datos de Hive Metastore o AWS Glue con mayor facilidad mediante los ajustes de configura ción. Añada una `hive.enabled: true` o varias `glue.enabled: true` `flink-conf` clasificaciones durante la creación del clúster para configurar automáticamente el catálogo de

Tipo	Descripción
	datos correspondiente. Este enfoque simplific ado elimina los pasos de configuración manual que antes eran necesarios.

Amazon EMR 7.8.0 - Notas de la versión de Flink

Configuración: EMR Flink funciona de forma inmediata con el S3A en todas las regiones y particiones. AWS

Amazon EMR 7.7.0 - Notas de la versión de Flink

- El shell SQL de Flink se puede invocar fácilmente con el comando `flink-sql-client` que está enlazado simbólicamente a `/usr/lib/flink/bin/sql-client.sh`

Amazon EMR 7.6.0 - Notas de la versión de Flink

Amazon EMR 7.6.0 - Características de Flink

- No hay cambios en la versión.

Amazon EMR 7.5.0 - Notas de la versión de Flink

Tipo	Descripción
Característica	Se agregó soporte para ejecutar tareas de Flink con un contenedor remoto.
Mejora	Haga que los subprocesos de exclusión e inclusión de vértices sean seguros.

Amazon EMR 7.5.0 - Características de Flink

- A partir de Amazon EMR 7.5.0, puede especificar una ubicación de Amazon S3 como ruta JAR cuando utilice los comandos CLI `run` y Apache `run-application` Flink. Al proporcionar una ruta S3, EMR descarga automáticamente el archivo JAR de Amazon S3 al almacenamiento de EBS

del clúster. Cada vez que especifique el mismo archivo JAR, EMR descarga la última versión de Amazon S3 en lugar de volver a utilizar el archivo JAR existente en el clúster.

- A partir de Amazon EMR 7.5.0, los clientes pueden pasar la ruta remota (una ubicación S3) como la ruta JAR con los comandos CLI de `run` Flink`run-application`. A continuación, el JAR se extrae automáticamente del almacén de S3 al almacenamiento de EBS del clúster. Si se vuelve a proporcionar el mismo JAR, se descarga el último de S3 y no se reutiliza el JAR existente en el clúster.

Amazon EMR 7.4.0 - Notas de la versión de Flink

Tipo	Descripción
Upgrade	La versión de Flink se ha actualizado a la 1.19.1.

Amazon EMR 7.3.0: notas de la versión de Flink

- De forma predeterminada, los clústeres con el cifrado en tránsito habilitado mediante la configuración de seguridad utilizarán TLS 1.3 para la comunicación interna entre los procesos de Flink, el punto final REST de Job Manager y el servidor de historial de trabajos de Flink.

Amazon EMR 7.2.0: notas de la versión de Flink

Tipo	Descripción
Mejoras	La asistencia permite añadir etiquetas personalizadas al servicio de Kubernetes por trabajo de Flink mediante la configuración `kubernetes.service.labels` .

Ganglia

> **Note**
>
> La última versión de Amazon EMR que incluye Ganglia es Amazon EMR 6.15.0. Para monitorizar el clúster, las versiones posteriores a la 6.15.0 incluyen el agente de Amazon CloudWatch .

El proyecto de código abierto Ganglia es un sistema distribuido y escalable diseñado para monitorizar clústeres y redes al mismo tiempo que minimiza el impacto en su rendimiento. Al habilitar Ganglia en su clúster, puede generar informes y ver el rendimiento del clúster en su conjunto, además de inspeccionar el rendimiento de cada una de las instancias de nodo individuales. Ganglia también está configurado para adquirir y visualizar las métricas de Hadoop y Spark. Para obtener más información acerca del proyecto de código abierto Ganglia, visite http://ganglia.info/.

Cuando se consulta la IU web de Ganglia en un navegador, verá una descripción general del rendimiento del clúster, con gráficos donde se detalla la carga, el uso de memoria, el uso de CPU y el tráfico de red del clúster. Debajo de las estadísticas del clúster hay gráficos para cada servidor individual en el clúster.

En la tabla siguiente, se muestra la versión de Ganglia incluida en la última versión de la serie 6.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Ganglia.

Para ver la versión de los componentes instalados con Ganglia en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Ganglia para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.15.0	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

En la tabla siguiente, se muestra la versión de Ganglia incluida en la última versión de la serie 5.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Ganglia.

Para ver la versión de los componentes instalados con Ganglia en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Ganglia para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.36.2	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Temas

- Creación de un clúster con Ganglia
- Ver métricas de Ganglia
- Métricas de Hadoop y Spark en Ganglia
- Historial de versiones de Ganglia

Creación de un clúster con Ganglia

> **Note**
>
> La última versión de Amazon EMR que incluye Ganglia es Amazon EMR 6.15.0. Para monitorizar el clúster, las versiones posteriores a la 6.15.0 incluyen el agente de Amazon CloudWatch .

Para crear un clúster con Ganglia utilizando la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Create cluster.
3. En Software configuration (Configuración de software), elija All Applications (Todas las aplicaciones), Core Hadoop o bien Spark.
4. Continúe con la creación del clúster con las configuraciones según proceda.

Para añadir Ganglia a un clúster mediante el AWS CLI

En el AWS CLI, puede añadir Ganglia a un clúster utilizando el `create-cluster --applications` parámetro. Si especifica solo Ganglia utilizando el parámetro `--applications`, Ganglia es la única aplicación instalada.

- Escriba el siguiente comando para añadir Ganglia al crear un clúster y *myKey* sustitúyalo por el nombre del EC2 key pair.

  > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Spark cluster with Ganglia" --release-label
 emr-6.15.0; \
```

```
--applications Name=Spark Name=Ganglia \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3 --use-default-roles
```

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos secundarios. Todos los nodos utilizan el tipo de instancia especificado en el comando.

> **Note**
>
> Si no ha creado anteriormente el rol de servicio EMR y el perfil de EC2 instancia predeterminados, escriba `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Ver métricas de Ganglia

> **Note**
>
> La última versión de Amazon EMR que incluye Ganglia es Amazon EMR 6.15.0. Para monitorizar el clúster, las versiones posteriores a la 6.15.0 incluyen el agente de Amazon CloudWatch .

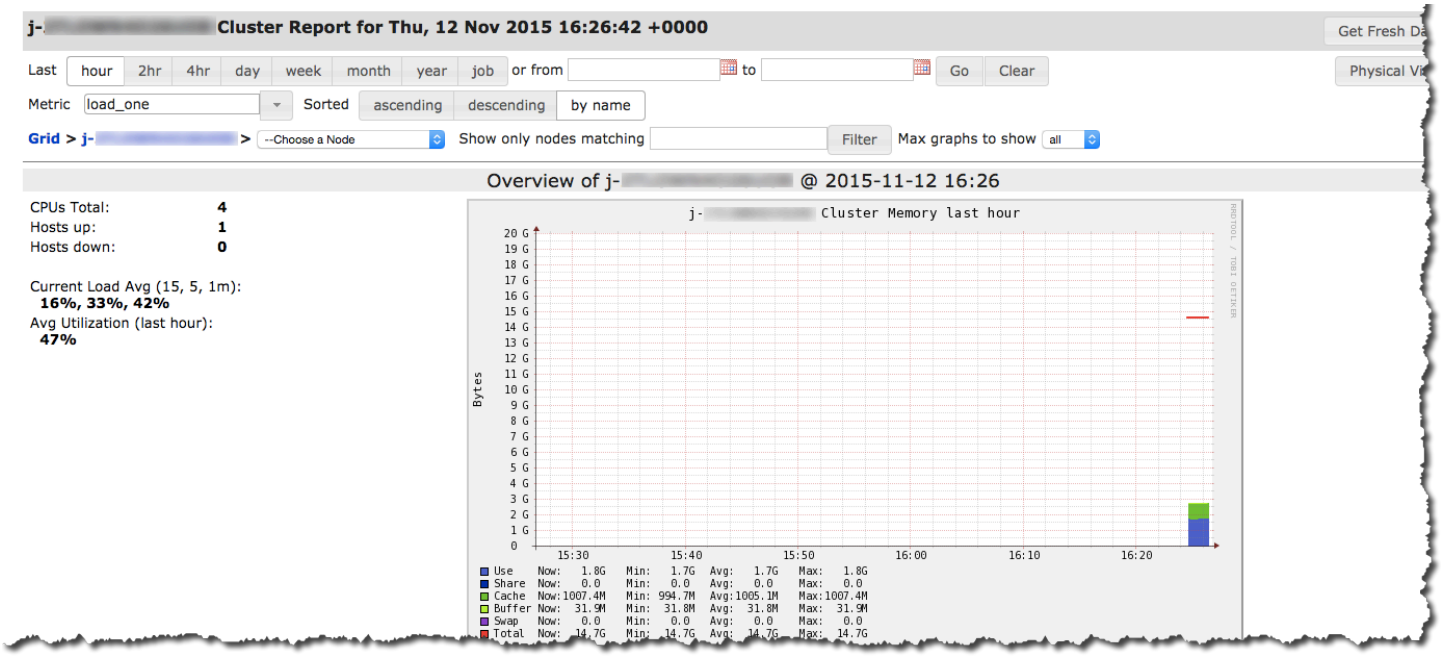
Ganglia proporciona una interfaz de usuario basada en web que puede utilizar para consultar las métricas que recopila Ganglia. Cuando ejecuta Ganglia en Amazon EMR, la interfaz web se ejecuta en el nodo maestro y se puede visualizar mediante el reenvío de puertos, también conocido como la creación de un túnel SSH. Para obtener más información sobre la visualización de interfaces web en Amazon EMR, consulte Ver interfaces web alojadas en clústeres de EMR en la Guía de administración de Amazon EMR.

Para ver la interfaz web de Ganglia

1. Utilice SSH para crear un túnel en el nodo principal y crear una conexión segura. Para obtener información sobre cómo crear un túnel SSH en el nodo maestro, consulte Opción 2, parte 1:

configurar un túnel SSH hacia el nodo maestro mediante el reenvío dinámico de puertos en la Guía de administración de Amazon EMR.

2. Instale un navegador web con una herramienta de proxy, como el FoxyProxy complemento para Firefox, para crear un proxy SOCKS para dominios del tipo *ec2*.amazonaws.com*. Para obtener más información, consulte Opción 2, parte 2: configurar los ajustes del proxy para ver los sitios web alojados en el nodo maestro en la Guía de administración de Amazon EMR.
3. Con el proxy configurado y la conexión SSH abierta, puede ver la interfaz de usuario de Ganglia abriendo una ventana del navegador con http://*master-public-dns-name*/ganglia/, donde *master-public-dns-name* se encuentra la dirección DNS pública del servidor maestro del clúster EMR.



Métricas de Hadoop y Spark en Ganglia

> **Note**
>
> La última versión de Amazon EMR que incluye Ganglia es Amazon EMR 6.15.0. Para monitorizar el clúster, las versiones posteriores a la 6.15.0 incluyen el agente de Amazon CloudWatch .

Ganglia notifica métricas de Hadoop para cada instancia. Los distintos tipos de métricas vienen prefijados por categoría: sistema de archivos distribuido (dfs.*), máquina virtual Java (jvm.*), (mapred.*) y llamadas a procedimientos remotos MapReduce (rpc.*).

Las métricas de Ganglia basadas en YARN como, por ejemplo, Spark y Hadoop no están disponibles para las versiones 4.4.0 y 4.5.0 de EMR. Utilice una versión posterior para usar estas métricas.

Las métricas de Ganglia para Spark suelen tener prefijos para el ID de la aplicación YARN y para Spark. DAGScheduler Por lo tanto, los prefijos siguen este formato:

- DAGScheduler.*
- application_xxxxxxxxxx_xxxx.driver.*
- application_xxxxxxxxxx_xxxx.executor.*

Historial de versiones de Ganglia

> **Note**
>
> La última versión de Amazon EMR que incluye Ganglia es Amazon EMR 6.15.0. Para monitorizar el clúster, las versiones posteriores a la 6.15.0 incluyen el agente de Amazon CloudWatch .

En la siguiente tabla, se muestra la versión de Ganglia incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Ganglia

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.36.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode,

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
		hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.36.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.36.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.35.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-6.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.34.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.33.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.33.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.31.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.31.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.30.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.29.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.28.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.28.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.27.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.27.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.26.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.25.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.24.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.24.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.23.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.23.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.22.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.21.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.20.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.20.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.18.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.18.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.17.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.16.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.16.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.14.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.13.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.12.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.11.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.5.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.5.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.3.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.2.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.0.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-5.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.6	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-4.9.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-4.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-4.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-4.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-4.7.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-4.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Etiqueta de versión de Amazon EMR	Versión de Ganglia	Componentes instalados con Ganglia
emr-4.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.2.0	3.6.0	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Apache Hadoop

Apache Hadoop es un marco de trabajo de software de Java de código abierto que admite el procesamiento de grandes conjuntos de datos a través de un clúster de instancias. Se puede ejecutar en una única instancia o en miles de instancias. Hadoop usa muchos modelos de procesamiento, como MapReduce el Tez, para distribuir el procesamiento en varias instancias y también usa un sistema de archivos distribuido llamado HDFS para almacenar datos en varias instancias. Hadoop monitoriza el estado de las instancias del clúster y puede recuperarse de errores en uno o varios nodos. De esta manera, Hadoop proporciona mayor capacidad de almacenamiento y procesamiento, así como mayor disponibilidad. Para obtener más información, consulte la documentación de Hadoop.

En la siguiente tabla, se muestra la versión de Hadoop incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Hadoop.

Para ver la versión de los componentes instalados con Hadoop en esta versión, consulte Versiones de componentes de la versión 7.9.0.

Información sobre la versión de Hadoop para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-7.9.0	Hadoop 3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

En la tabla siguiente, se muestra la versión de Hadoop incluida en la última versión de la serie 6.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Hadoop.

Para ver la versión de los componentes instalados con Hadoop en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Hadoop para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.15.0	Hadoop 3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

En la tabla siguiente, se muestra la versión de Hadoop incluida en la última versión de la serie 5.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Hadoop.

Para ver la versión de los componentes instalados con Hadoop en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Hadoop para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.36.2	Hadoop 2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
		nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server

A partir de Amazon EMR 5.18.0, puede utilizar el repositorio de artefactos de Amazon EMR para compilar el código de trabajos con las versiones exactas de las bibliotecas y dependencias que están disponibles con determinadas versiones de Amazon EMR. Para obtener más información, consulte Comprobación de dependencias mediante el repositorio de artefactos de Amazon EMR.

Temas

- Configurar Hadoop
- Cifrado transparente en el HDFS en Amazon EMR
- Crear o ejecutar una aplicación de Hadoop
- Leer objetos restaurados
- Activar el reconocimiento de acceso no uniforme a la memoria para los contenedores YARN
- Embalaje de contenedores YARN
- Historial de versiones de Hadoop

Configurar Hadoop

En las secciones siguientes se ofrecen las opciones de configuración predeterminadas para daemons de Hadoop, tareas y HDFS.

Temas

- Configuración de tareas
- Ajustes de configuración del daemon de Hadoop
- Configuración de HDFS

Configuración de tareas

Puede establecer variables de configuración para ajustar el rendimiento de sus MapReduce trabajos. En esta sección se proporcionan los valores predeterminados de los ajustes importantes. Los valores predeterminados varían en función del tipo de EC2 instancia del nodo utilizado en el clúster. HBase está disponible cuando se utiliza la versión 4.6.0 y versiones posteriores de Amazon EMR. Cuando HBase se instala, se utilizan diferentes valores predeterminados. Estos valores se proporcionan junto con los valores predeterminados iniciales.

Hadoop 2 usa dos parámetros, uno `mapreduce.map.java.opts` y uno para configurar la memoria para mapear y reducir`mapreduce.reduce.java.opts`, respectivamente. JVMs Estos parámetros sustituyen a la opción de configuración única `mapreduce.map.java.opts` de versiones de Hadoop anteriores.

Del mismo modo, `mapred.job.jvm.num.tasks` sustituye a `mapred.job.reuse.jvm.num.tasks` en Hadoop 2.7.2 y versiones posteriores. Amazon EMR establece este valor en 20, independientemente del tipo de EC2 instancia. Puede invalidar este valor con la clasificación de configuración `mapred-site`. La configuración de un valor de `-1` indica que una JVM se puede volver a utilizar para un número infinito de tareas dentro de una solo trabajo, y un valor de 1 indica que se genera una nueva JVM para cada tarea.

Por ejemplo, para establecer el valor de `mapred.job.jvm.num.tasks` en -1 puede crear un archivo con el siguiente contenido:

```
[
 {
  "Classification": "mapred-site",
  "Properties": {
   "mapred.job.jvm.num.tasks": "-1"
  }
 }
]
```

Cuando utilice el `create-cluster` comando o el `modify-instance-groups` comando de AWS CLI, podrá hacer referencia al archivo de configuración JSON. En el siguiente ejemplo, el archivo de configuración se guarda como `myConfig.json` y se almacena en Amazon S3.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--instance-count 3 --applications Name=Hadoop --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json \
--use-default-roles
```

Para cambiar los valores predeterminados que se enumeran a continuación, utilice la clasificación de configuración `mapred-site` de la misma forma, y establecer varias clasificaciones de configuración mediante un único archivo JSON. Para obtener más información, consulte Configuración de aplicaciones.

Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Valores predeterminados para los ajustes de configuración de tareas

> **Note**
>
> Los valores especificados para cada instancia están de acuerdo con la última versión de EMR. Las versiones anteriores pueden tener diferentes configuraciones predeterminadas.

Tipos de instancias

- Instancias c1
- Instancias c3
- Instancias c4
- Instancias c5

- Instancias c5a
- Instancias c5ad
- Instancias c5d
- Instancias c5n
- Instancias c6a
- Instancias c6g
- Instancias c6gd
- Instancias c6gn
- Instancias c6i
- Instancias C6id
- Instancias c6in
- Instancias C7a
- Instancias c7g
- Instancias c7gd
- Instancias C7gn
- Instancias c7i
- instancias c7i-flex
- instancias c8g
- instancias c8gd
- Instancias d2
- Instancias d3
- Instancias d3en
- instancias f2
- Instancias g3
- Instancias g3s
- Instancias g4dn
- Instancias g5
- 6 instancias

- 6 instancias
- instancias gr6
- Instancias h1
- Instancias i2
- Instancias i3
- Instancias i3en
- Instancias i4g
- Instancias i4i
- Instancias i7i
- Instancias i7ie
- instancias i8g
- Instancias im4gn
- Instancias is4gen
- Instancias m1
- Instancias m2
- Instancias m3
- Instancias m4
- Instancias m5
- Instancias m5a
- Instancias m5ad
- Instancias m5d
- Instancias m5dn
- Instancias m5n
- Instancias m5zn
- Instancias m6a
- Instancias m6g
- Instancias m6gd
- Instancias m6i

- Instancias m6id
- Instancias m6idn
- Instancias m6in
- Instancias m7a
- Instancias m7g
- Instancias m7gd
- Instancias m7i
- Instancias m7i-flex
- instancias m8g
- instancias m8gd
- Instancias p2
- Instancias p3
- instancias p4d
- Instancias p5
- Instancias r3
- Instancias r4
- Instancias r5
- Instancias r5a
- Instancias r5ad
- Instancias r5b
- Instancias r5d
- Instancias r5dn
- Instancias r5n
- Instancias r6a
- Instancias r6g
- Instancias r6gd
- Instancias r6i
- Instancias r6id
- Instancias r6idn

- Instancias r6in
- Instancias r7a
- Instancias r7g
- Instancias r7gd
- Instancias r7i
- Instancias r7iz
- instancias r8g
- instancias r8gd
- Instancias x1
- Instancias x1e
- Instancias x2gd
- Instancias x2idn
- Instancias x2iedn
- instancias x8g
- Instancias z1d

Instancias c1

c1.medium

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx288m	-Xmx288m
mapreduce.java.opts	-Xmx288m	-Xmx288m
mapreduce.map.memory.mb	512	512
mapreduce.reduce.memory.mb	512	512
yarn.app.mapreduce.am.resource.mb	512	512
yarn.scheduler.minimum-allocation-mb	1	1

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1024	512

c1.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1024	1024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2048	2560
yarn.nodemanager.resource.memory-mb	5120	2560

Instancias c3

c3.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c3.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2 880	2 880
yarn.app.mapreduce.am.resource.mb	2 880	2 880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.nodemanager.resource.memory-mb	11520	5760

c3.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2 880	2 880
yarn.app.mapreduce.am.resource.mb	2 880	2 880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c3.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1331m	-Xmx1331m
mapreduce.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.map.memory.mb	1664	1664

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.reduce.memory.mb	3328	3328
yarn.app.mapreduce.am.resource.mb	3328	3328
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

Instancias c4

c4.large

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx717m	-Xmx717m
mapreduce.java.opts	-Xmx1434m	-Xmx1434m
mapreduce.map.memory.mb	896	896
mapreduce.reduce.memory.mb	1792	1792
yarn.app.mapreduce.am.resource.mb	1792	1792
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1792	896

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.nodemanager.resource.memory-mb	1792	896

c4.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c4.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.reduce.memory.mb	2 880	2 880
yarn.app.mapreduce.am.resource.mb	2 880	2 880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

c4.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2 880	2 880
yarn.app.mapreduce.am.resource.mb	2 880	2 880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c4.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1183m	-Xmx1183m
mapreduce.java.opts	-Xmx2366m	-Xmx2366m
mapreduce.map.memory.mb	1479	1479
mapreduce.reduce.memory.mb	2958	2958
yarn.app.mapreduce.am.resource.mb	2958	2958
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

Instancias c5

c5.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

c5.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5.9xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5.18xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1547m	-Xmx1547m
mapreduce.java.opts	-Xmx3094m	-Xmx3094m
mapreduce.map.memory.mb	1934	1934
mapreduce.reduce.memory.mb	3868	3868
yarn.app.mapreduce.am.resource.mb	3868	3868
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

Instancias c5a

c5a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c5a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5a.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c5a.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c5a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5a.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

Instancias c5ad

c5ad.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c5ad.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5ad.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c5ad.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c5ad.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c5ad.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5ad.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

Instancias c5d

c5d.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

c5d.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5d.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5d.9xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5d.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1502m	-Xmx1502m
mapreduce.java.opts	-Xmx3004m	-Xmx3004m
mapreduce.map.memory.mb	1877	1877
mapreduce.reduce.memory.mb	3754	3754
yarn.app.mapreduce.am.resource.mb	3754	3754
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5d.18xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1547m	-Xmx1547m
mapreduce.java.opts	-Xmx3094m	-Xmx3094m
mapreduce.map.memory.mb	1934	1934
mapreduce.reduce.memory.mb	3868	3868
yarn.app.mapreduce.am.resource.mb	3868	3868
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5d.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

Instancias c5n

c5n.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1613m	-Xmx1613m
mapreduce.java.opts	-Xmx3226m	-Xmx3226m
mapreduce.map.memory.mb	2016	2016
mapreduce.reduce.memory.mb	4032	4032
yarn.app.mapreduce.am.resource.mb	4032	4032
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8064	4032
yarn.nodemanager.resource.memory-mb	8064	4032

c5n.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1613m	-Xmx1613m
mapreduce.java.opts	-Xmx3226m	-Xmx3226m
mapreduce.map.memory.mb	2016	2016
mapreduce.reduce.memory.mb	4032	4032
yarn.app.mapreduce.am.resource.mb	4032	4032
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	16128	8064
yarn.nodemanager.resource.memory-mb	16128	8064

c5n.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1741m	-Xmx1741m
mapreduce.java.opts	-Xmx3482m	-Xmx3482m
mapreduce.map.memory.mb	2176	2176
mapreduce.reduce.memory.mb	4352	4352
yarn.app.mapreduce.am.resource.mb	4352	4352

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	34816	17408
yarn.nodemanager.resource.memory-mb	34816	17408

c5n.9xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2002m	-Xmx2002m
mapreduce.java.opts	-Xmx4004m	-Xmx4004m
mapreduce.map.memory.mb	2503	2503
mapreduce.reduce.memory.mb	5006	5006
yarn.app.mapreduce.am.resource.mb	5006	5006
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

c5n.18xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2094m	-Xmx2094m
mapreduce.java.opts	-Xmx4188m	-Xmx4188m
mapreduce.map.memory.mb	2617	2617
mapreduce.reduce.memory.mb	5234	5234
yarn.app.mapreduce.am.resource.mb	5234	5234
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

Instancias c6a

c6a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6a.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6a.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6a.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6a.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c6a.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

Instancias c6g

c6g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6g.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instancias c6gd

c6gd.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6gd.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gd.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instancias c6gn

c6gn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6gn.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gn.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gn.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instancias c6i

c6i.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6i.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6i.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6i.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6i.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias C6id

c6id.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6id.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6id.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6id.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6id.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6id.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6id.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6id.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias c6in

c6in.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6in.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6in.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6in.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6in.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6in.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6in.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6in.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias C7a

c7a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7a.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7a.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7a.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7a.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c7a.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

Instancias c7g

c7g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7g.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instancias c7gd

c7gd.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7gd.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gd.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instancias C7gn

c7gn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7gn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7gn.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gn.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gn.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7gn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

Instancias c7i

c7i.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7i.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7i.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7i.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instancias c7i-flex

c7i-flex.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7i-flex.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7i-flex.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i-flex.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i-flex.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7i-flex.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

instancias c8g

c8g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c8g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c8g.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c8g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c8g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c8g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8g.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c8g.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

instancias c8gd

c8gd.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c8gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c8gd.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c8gd.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c8gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c8gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8gd.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c8gd.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

Instancias d2

d2.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d2.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d2.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d2.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2417m	-Xmx2417m
mapreduce.java.opts	-Xmx4834m	-Xmx4834m
mapreduce.map.memory.mb	3021	3021
mapreduce.reduce.memory.mb	6042	6042
yarn.app.mapreduce.am.resource.mb	6042	6042

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30194
yarn.nodemanager.resource.memory-mb	241664	30194

Instancias d3

d3.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias d3en

d3en.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

d3en.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3en.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3en.6xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

d3en.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3en.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

instancias f2

f2.6xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMX8055M	-XMX8055 m
mapreduce.java.opts	-XMX 16110 m	-XMX 16110 m
mapreduce.map.memory.mb	10069	10069
mapreduce.reduce.memory.mb	20138	20138
yarn.app.mapreduce.am.resource.mb	20138	20138
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	20146
yarn.nodemanager.resource.memory-mb	241664	20146

f2.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx8192m	-XMX8192 m
mapreduce.java.opts	-XMX 16384 m	-XMX16384 m
mapreduce.map.memory.mb	10240	10240
mapreduce.reduce.memory.mb	20480	20480
yarn.app.mapreduce.am.resource.mb	20480	20480

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	20480
yarn.nodemanager.resource.memory-mb	491520	20480

f2.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx8294m	-Xmx8294m
mapreduce.java.opts	-Xmx16588m	-Xmx16588m
mapreduce.map.memory.mb	10368	10368
mapreduce.reduce.memory.mb	20736	20736
yarn.app.mapreduce.am.resource.mb	20736	20736
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	20736
yarn.nodemanager.resource.memory-mb	1990656	20736

Instancias g3

g3.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g3.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g3.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias g3s

g3s.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

Instancias g4dn

g4dn.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

g4dn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

g4dn.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

g4dn.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

g4dn.12xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

g4dn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

Instancias g5

g5.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

g5.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g5.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g5.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g5.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

g5.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g5.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

g5.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

6 instancias

g6.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2 880	2 880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

g6.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2 880	2 880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

g6.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.java.opts	-XMx5324m	-XMx5324 m
mapreduce.map.memory.mb	3328	3328
mapreduce.reduce.memory.mb	6656	6656
yarn.app.mapreduce.am.resource.mb	6656	6656
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

g6.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx2893m	-XMx2893 m
mapreduce.java.opts	-XMX 5786 m	-XMX 5786 m
mapreduce.map.memory.mb	3616	3616
mapreduce.reduce.memory.mb	7232	7232
yarn.app.mapreduce.am.resource.mb	7232	7232
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

g6.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx2953m	-XMX 2953 m
mapreduce.java.opts	-XMX 5906 m	-XMX 5906 m
mapreduce.map.memory.mb	3691	3691
mapreduce.reduce.memory.mb	7382	7382
yarn.app.mapreduce.am.resource.mb	7382	7382

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	177152	29512
yarn.nodemanager.resource.memory-mb	177152	29512

g6.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx2995m	-XMX 295 m
mapreduce.java.opts	-XMX 5990 m	-XMX 5990 m
mapreduce.map.memory.mb	3744	3744
mapreduce.reduce.memory.mb	7488	7488
yarn.app.mapreduce.am.resource.mb	7488	7488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	239616	29952
yarn.nodemanager.resource.memory-mb	239616	29952

g6.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx3038m	-XMX 3038 m
mapreduce.java.opts	-XMX 6076 m	-XMX 6076 m
mapreduce.map.memory.mb	3797	3797
mapreduce.reduce.memory.mb	7594	7594
yarn.app.mapreduce.am.resource.mb	7594	7594
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	364544	30408
yarn.nodemanager.resource.memory-mb	364544	30408

g6.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	737280	30720
yarn.nodemanager.resource.memory-mb	737280	30720

6 instancias

g6e.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g6e.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g6e.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g6g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g6e.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

g6e.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

g6e.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

g6e.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instancias gr6

gr6.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx5786m	-XMX 5786 m
mapreduce.java.opts	-XMX 11572 m	-XMX 11572 m
mapreduce.map.memory.mb	7232	7232
mapreduce.reduce.memory.mb	14464	14464
yarn.app.mapreduce.am.resource.mb	14464	14464

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

gr6.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx5990m	-XMX 5990 m
mapreduce.java.opts	-XMX 11980 m	-XMX 11980 m
mapreduce.map.memory.mb	7488	7488
mapreduce.reduce.memory.mb	14976	14976
yarn.app.mapreduce.am.resource.mb	14976	14976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	239616	29952
yarn.nodemanager.resource.memory-mb	239616	29952

Instancias h1

h1.2xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

h1.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

h1.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

h1.16xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

Instancias i2

i2.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i2.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i2.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i2.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias i3

i3.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i3.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i3.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i3.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i3.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias i3en

i3en.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

i3en.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

i3en.3xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

i3en.6xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

i3en.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

i3en.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instancias i4g

i4g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i4g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i4g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias i4i

i4i.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i4i.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i4i.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i4i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i4i.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i4i.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instancias i7i

i7i.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i7i.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i7i.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i7i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i7i.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7i.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instancias i7ie

i7ie.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7ie.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i7ie.3xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.java.opts	-XMx 11400 m	-XMX 11400 m
mapreduce.map.memory.mb	7125	7125
mapreduce.reduce.memory.mb	14250	14250
yarn.app.mapreduce.am.resource.mb	14250	14250
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28504
yarn.nodemanager.resource.memory-mb	85504	28504

i7ie.6xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMX5974m	-XMX 5974 m
mapreduce.java.opts	-XMX 11948 m	-XMX 11948 m
mapreduce.map.memory.mb	7467	7467
mapreduce.reduce.memory.mb	14934	14934
yarn.app.mapreduce.am.resource.mb	14934	14934

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29860
yarn.nodemanager.resource.memory-mb	179200	29860

i7ie.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7ie.18xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx6155m	-XMX 6155 m
mapreduce.java.opts	-XMx 12310 m	-XMx12310 m
mapreduce.map.memory.mb	7694	7694
mapreduce.reduce.memory.mb	15388	15388
yarn.app.mapreduce.am.resource.mb	15388	15388
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	553984	30792
yarn.nodemanager.resource.memory-mb	553984	30792

i7ie.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7ie.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instancias i8g

i8g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i8g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i8g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i8g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i8g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i8g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i8g.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i8g.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instancias im4gn

im4gn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

im4gn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

im4gn.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

im4gn.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

im4gn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias is4gen

is4gen.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3514m	-Xmx3514m
mapreduce.java.opts	-Xmx7028m	-Xmx7028m
mapreduce.map.memory.mb	4393	4393
mapreduce.reduce.memory.mb	8786	8786
yarn.app.mapreduce.am.resource.mb	8786	8786

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	17572,12	8786,06
yarn.nodemanager.resource.memory-mb	17572,12	8786,06

is4gen.2xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3866m	-Xmx3866m
mapreduce.java.opts	-Xmx7732m	-Xmx7732m
mapreduce.map.memory.mb	4832	4832
mapreduce.reduce.memory.mb	9664	9664
yarn.app.mapreduce.am.resource.mb	9664	9664
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

is4gen.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4275m	-Xmx4275m
mapreduce.java.opts	-Xmx8550m	-Xmx8550m
mapreduce.map.memory.mb	5344	5344
mapreduce.reduce.memory.mb	10688	10688
yarn.app.mapreduce.am.resource.mb	10688	10688
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32064
yarn.nodemanager.resource.memory-mb	85504	32064

is4gen.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4480m	-Xmx4480m
mapreduce.java.opts	-Xmx8960m	-Xmx8960m
mapreduce.map.memory.mb	5600	5600
mapreduce.reduce.memory.mb	11200	11200
yarn.app.mapreduce.am.resource.mb	11200	11200

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	22400
yarn.nodemanager.resource.memory-mb	179200	22400

Instancias m1

m1.small

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx288m	-Xmx288m
mapreduce.java.opts	-Xmx288m	-Xmx288m
mapreduce.map.memory.mb	512	512
mapreduce.reduce.memory.mb	512	512
yarn.app.mapreduce.am.resource.mb	512	512
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1024	512

m1.medium

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx768m	-Xmx768m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1024	1024
yarn.app.mapreduce.am.resource.mb	1024	1024
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2048	1024
yarn.nodemanager.resource.memory-mb	2048	1024

m1.large

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1024m	-Xmx1024m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	1536	1536
yarn.app.mapreduce.am.resource.mb	1536	1536

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3072	2560
yarn.nodemanager.resource.memory-mb	5120	2560

m1.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	6144
yarn.nodemanager.resource.memory-mb	12288	6144

Instancias m2

m2.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1024	1024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	7168	7168
yarn.nodemanager.resource.memory-mb	14336	7168

m2.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	15360
yarn.nodemanager.resource.memory-mb	30720	15360

m2.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	30720
yarn.nodemanager.resource.memory-mb	61440	30720

Instancias m3

m3.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2 880	2 880
yarn.app.mapreduce.am.resource.mb	2 880	2 880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m3.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2 880	2 880
yarn.app.mapreduce.am.resource.mb	2 880	2 880

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

Instancias m4

m4.large

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

m4.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m4.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m4.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1434m	-Xmx1434m
mapreduce.java.opts	-Xmx2868m	-Xmx2868m
mapreduce.map.memory.mb	1792	1792
mapreduce.reduce.memory.mb	3584	3584
yarn.app.mapreduce.am.resource.mb	3584	3584
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m4.10xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx1557m	-Xmx1557m
mapreduce.java.opts	-Xmx3114m	-Xmx3114m
mapreduce.map.memory.mb	1946	1946
mapreduce.reduce.memory.mb	3892	3892
yarn.app.mapreduce.am.resource.mb	3892	3892
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	155648	31104
yarn.nodemanager.resource.memory-mb	155648	31104

m4.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx1587m	-Xmx1587m
mapreduce.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.map.memory.mb	1984	1984
mapreduce.reduce.memory.mb	3968	3968
yarn.app.mapreduce.am.resource.mb	3968	3968

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

Instancias m5

m5.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5.12xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

Instancias m5a

m5a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5a.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5a.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5a.12xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5a.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

Instancias m5ad

m5ad.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5ad.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5ad.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5ad.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5ad.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5ad.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5ad.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

Instancias m5d

m5d.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5d.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5d.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5d.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5d.12xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5d.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5d.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

Instancias m5dn

m5dn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5dn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5dn.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5dn.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5dn.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5dn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5dn.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

Instancias m5n

m5n.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5n.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5n.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5n.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5n.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5n.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5n.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

Instancias m5zn

m5zn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2 880	2 880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m5zn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5zn.3xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2577m	-Xmx2577m
mapreduce.java.opts	-Xmx5154m	-Xmx5154m
mapreduce.map.memory.mb	3221	3221
mapreduce.reduce.memory.mb	6442	6442
yarn.app.mapreduce.am.resource.mb	6442	6442

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

m5zn.6xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

m5zn.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

Instancias m6a

m6a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6a.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6a.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6a.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6a.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m6a.48xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instancias m6g

m6g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6g.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias m6gd

m6gd.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6gd.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6gd.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias m6i

m6i.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6i.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6i.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6i.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6i.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias m6id

m6id.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6id.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6id.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6id.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6id.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6id.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6id.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6id.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias m6idn

m6idn.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6idn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6idn.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6idn.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6idn.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6idn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6idn.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6idn.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias m6in

m6in.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6in.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6in.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6in.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6in.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6in.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6in.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6in.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias m7a

m7a.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7a.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7a.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m7a.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7a.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m7a.48xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instancias m7g

m7g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7g.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias m7gd

m7gd.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7gd.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7gd.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias m7i

m7i.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7i.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m7i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m7i.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7i.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instancias m7i-flex

m7i-flex.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7i-flex.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i-flex.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7i-flex.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7i-flex.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7i-flex.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

instancias m8g

m8g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m8g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m8g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m8g.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m8g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m8g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m8g.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m8g.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

instancias m8gd

m8gd.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m8gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m8gd.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m8gd.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m8gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m8gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m8gd.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m8gd.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

Instancias p2

p2.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p2.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

p2.16xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx9267m	-Xmx9267m
mapreduce.java.opts	-Xmx18534m	-Xmx18534m
mapreduce.map.memory.mb	11584	11584
mapreduce.reduce.memory.mb	23168	23168
yarn.app.mapreduce.am.resource.mb	23168	23168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	23168
yarn.nodemanager.resource.memory-mb	741376	23168

Instancias p3

p3.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p3.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

p3.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

instancias p4d

p4d.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx 9302 m	-XMx 9302 m
mapreduce.java.opts	-XMX 18604 m	-XMx18604 m
mapreduce.map.memory.mb	11627	11627
mapreduce.reduce.memory.mb	23254	23254
yarn.app.mapreduce.am.resource.mb	23254	23254
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1116160	23222
yarn.nodemanager.resource.memory-mb	1116160	23222

Instancias p5

p5.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx8294m	-Xmx8294m
mapreduce.java.opts	-Xmx16588m	-Xmx16588m
mapreduce.map.memory.mb	10368	10368
mapreduce.reduce.memory.mb	20736	20736

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.app.mapreduce.am.resource.mb	20736	20736
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	20736
yarn.nodemanager.resource.memory-mb	1990656	20736

Instancias r3

r3.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r3.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r3.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r3.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

Instancias r4

r4.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r4.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r4.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r4.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r4.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias r5

r5.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instancias r5a

r5a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5a.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5a.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5a.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instancias r5ad

r5ad.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5ad.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5ad.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5ad.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5ad.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5ad.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6246m	-Xmx6246m
mapreduce.java.opts	-Xmx12492m	-Xmx12492m
mapreduce.map.memory.mb	7808	7808
mapreduce.reduce.memory.mb	15616	15616
yarn.app.mapreduce.am.resource.mb	15616	15616
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	499712	31232
yarn.nodemanager.resource.memory-mb	499712	31232

r5ad.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instancias r5b

r5b.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5b.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5b.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5b.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5b.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5b.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5b.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instancias r5d

r5d.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5d.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5d.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5d.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5d.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5d.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5d.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

Instancias r5dn

r5dn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5dn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5dn.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5dn.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5dn.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5dn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5dn.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instancias r5n

r5n.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5n.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5n.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5n.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5n.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5n.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5n.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

Instancias r6a

r6a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6a.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6a.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6a.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6a.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r6a.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instancias r6g

r6g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias r6gd

r6gd.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6gd.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6gd.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias r6i

r6i.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6i.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6i.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6i.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6i.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

Instancias r6id

r6id.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6id.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6id.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6id.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6id.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6id.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6id.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6id.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instancias r6idn

r6idn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6idn.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6idn.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6idn.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6idn.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6idn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6idn.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6idn.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instancias r6in

r6in.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6in.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6in.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6in.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6in.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6in.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6in.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6in.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instancias r7a

r7a.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7a.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7a.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7a.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7a.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7a.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7a.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7a.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r7a.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instancias r7g

r7g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias r7gd

r7gd.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7gd.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7gd.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

Instancias r7i

r7i.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7i.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7i.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7i.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7i.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7i.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7i.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7i.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instancias r7iz

r7iz.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7iz.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7iz.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7iz.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7iz.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7iz.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7iz.32xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

instancias r8g

r8g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r8g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r8g.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r8g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r8g.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r8g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r8g.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r8g.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

instancias r8gd

r8gd.xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r8gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r8gd.4xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r8gd.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r8gd.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r8gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r8gd.24xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r8gd.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

Instancias x1

x1.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.java.opts	-Xmx24116m	-Xmx24116m
mapreduce.map.memory.mb	15072	15072
mapreduce.reduce.memory.mb	30144	30144
yarn.app.mapreduce.am.resource.mb	30144	30144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

x1.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12109m	-Xmx12109m
mapreduce.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.map.memory.mb	15136	15136
mapreduce.reduce.memory.mb	30272	30272
yarn.app.mapreduce.am.resource.mb	30272	30272

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937408	30272
yarn.nodemanager.resource.memory-mb	1937408	30272

Instancias x1e

x1e.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx22682m	-Xmx22682m
mapreduce.java.opts	-Xmx45364m	-Xmx45364m
mapreduce.map.memory.mb	28352	28352
mapreduce.reduce.memory.mb	56704	56704
yarn.app.mapreduce.am.resource.mb	56704	56704
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	113408	0
yarn.nodemanager.resource.memory-mb	113408	0

x1e.2xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx23501m	-Xmx23501m
mapreduce.java.opts	-Xmx47002m	-Xmx47002m
mapreduce.map.memory.mb	29376	29376
mapreduce.reduce.memory.mb	58752	58752
yarn.app.mapreduce.am.resource.mb	58752	58752
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	235008	0
yarn.nodemanager.resource.memory-mb	235008	0

x1e.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx23910m	-Xmx23910m
mapreduce.java.opts	-Xmx47820m	-Xmx47820m
mapreduce.map.memory.mb	29888	29888
mapreduce.reduce.memory.mb	59776	59776
yarn.app.mapreduce.am.resource.mb	59776	59776

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	478208	0
yarn.nodemanager.resource.memory-mb	478208	0

x1e.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24115m	-Xmx24115m
mapreduce.java.opts	-Xmx48230m	-Xmx48230m
mapreduce.map.memory.mb	30144	30144
mapreduce.reduce.memory.mb	60288	60288
yarn.app.mapreduce.am.resource.mb	60288	60288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	0
yarn.nodemanager.resource.memory-mb	964608	0

x1e.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.java.opts	-Xmx48436m	-Xmx48436m
mapreduce.map.memory.mb	30272	30272
mapreduce.reduce.memory.mb	60544	60544
yarn.app.mapreduce.am.resource.mb	60544	60544
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937408	0
yarn.nodemanager.resource.memory-mb	1937408	0

x1e.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24269m	-Xmx24269m
mapreduce.java.opts	-Xmx48538m	-Xmx48538m
mapreduce.map.memory.mb	30336	30336
mapreduce.reduce.memory.mb	60672	60672
yarn.app.mapreduce.am.resource.mb	60672	60672

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3883008	0
yarn.nodemanager.resource.memory-mb	3883008	0

Instancias x2gd

x2gd.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

x2gd.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

x2gd.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

x2gd.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

x2gd.12xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

x2gd.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

Instancias x2idn

x2idn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

x2idn.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12425m	-Xmx12425m
mapreduce.java.opts	-Xmx24850m	-Xmx24850m
mapreduce.map.memory.mb	15531	15531
mapreduce.reduce.memory.mb	31062	31062
yarn.app.mapreduce.am.resource.mb	31062	31062
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

x2idn.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12442m	-Xmx12442m
mapreduce.java.opts	-Xmx24884m	-Xmx24884m
mapreduce.map.memory.mb	15552	15552
mapreduce.reduce.memory.mb	31104	31104
yarn.app.mapreduce.am.resource.mb	31104	31104

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	31104
yarn.nodemanager.resource.memory-mb	1990656	31104

Instancias x2iedn

x2iedn.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx23347m	-Xmx23347m
mapreduce.java.opts	-Xmx46694m	-Xmx46694m
mapreduce.map.memory.mb	29184	29184
mapreduce.reduce.memory.mb	58368	58368
yarn.app.mapreduce.am.resource.mb	58368	58368
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	0
yarn.nodemanager.resource.memory-mb	116736	0

x2iedn.2xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.java.opts	-Xmx48332m	-Xmx48332m
mapreduce.map.memory.mb	30208	30208
mapreduce.reduce.memory.mb	60416	60416
yarn.app.mapreduce.am.resource.mb	60416	60416
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	0
yarn.nodemanager.resource.memory-mb	241664	0

x2iedn.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.java.opts	-Xmx49152m	-Xmx49152m
mapreduce.map.memory.mb	30720	30720
mapreduce.reduce.memory.mb	61440	61440
yarn.app.mapreduce.am.resource.mb	61440	61440

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	0
yarn.nodemanager.resource.memory-mb	491520	0

x2iedn.8xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx24781m	-Xmx24781m
mapreduce.java.opts	-Xmx49562m	-Xmx49562m
mapreduce.map.memory.mb	30976	30976
mapreduce.reduce.memory.mb	61952	61952
yarn.app.mapreduce.am.resource.mb	61952	61952
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	0
yarn.nodemanager.resource.memory-mb	991232	0

x2iedn.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24883m	-Xmx24883m
mapreduce.java.opts	-Xmx49766m	-Xmx49766m
mapreduce.map.memory.mb	31104	31104
mapreduce.reduce.memory.mb	62208	62208
yarn.app.mapreduce.am.resource.mb	62208	62208
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	0
yarn.nodemanager.resource.memory-mb	1990656	0

x2iedn.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24918m	-Xmx24918m
mapreduce.java.opts	-Xmx49836m	-Xmx49836m
mapreduce.map.memory.mb	31147	31147
mapreduce.reduce.memory.mb	62294	62294
yarn.app.mapreduce.am.resource.mb	62294	62294

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	-32
yarn.nodemanager.resource.memory-mb	2990080	-32

x2iedn.32xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx24934m	-Xmx24934m
mapreduce.java.opts	-Xmx49868m	-Xmx49868m
mapreduce.map.memory.mb	31168	31168
mapreduce.reduce.memory.mb	62336	62336
yarn.app.mapreduce.am.resource.mb	62336	62336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3989504	0
yarn.nodemanager.resource.memory-mb	3989504	0

instancias x8g

x8g.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

x8g.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

x8g.4xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

x8g.8xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

x8g.12xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

x8g.16xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

x8g.24xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx12425m	-Xmx12425m
mapreduce.java.opts	-Xmx24850m	-Xmx24850m
mapreduce.map.memory.mb	15531	15531
mapreduce.reduce.memory.mb	31062	31062
yarn.app.mapreduce.am.resource.mb	31062	31062
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

x8g.48xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-XMx12458m	-XMx12458 m
mapreduce.java.opts	-XMX 24916 m	-XMX 24916 m
mapreduce.map.memory.mb	15573	15573
mapreduce.reduce.memory.mb	31146	31146
yarn.app.mapreduce.am.resource.mb	31146	31146

Opción de configuración	Valor predeterminado	Con instalado HBase
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	31210
yarn.nodemanager.resource.memory-mb	2990080	31210

Instancias z1d

z1d.xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

z1d.2xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

z1d.3xlarge

Opción de configuración	Valor predeterminado	Con HBase instalado
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018

Opción de configuración	Valor predeterminado	Con HBase instalado
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

z1d.6xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

z1d.12xlarge

Opción de configuración	Valor predeterminado	Con instalado HBase
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

Ajustes de configuración del daemon de Hadoop

Los ajustes del daemon de Hadoop son diferentes en función del tipo de EC2 instancia que utiliza un nodo del clúster. En las tablas siguientes se muestran los ajustes de configuración predeterminados para cada tipo de EC2 instancia.

Para personalizar estos ajustes, utilice la clasificación de configuración hadoop-env. Para obtener más información, consulte Configuración de aplicaciones.

Tipos de instancias

- Instancias c1
- Instancias c3
- Instancias c4
- Instancias c5

- Instancias c5a
- Instancias c5ad
- Instancias c5d
- Instancias c5n
- Instancias c6a
- Instancias c6g
- Instancias c6gd
- Instancias c6gn
- Instancias c6i
- Instancias C6id
- Instancias c6in
- Instancias c7a
- Instancias c7g
- Instancias c7gd
- Instancias C7gn
- Instancias c7i
- Instancias c7i-flex
- Instancias c8g
- instancias c8gd
- Instancias d2
- Instancias d3
- Instancias d3en
- instancias f2
- Instancias g3
- Instancias g3s
- Instancias g4dn
- Instancias g5
- Instancias g6
- Instancias g6e
- Instancias gr6

- Instancias h1
- Instancias i2
- Instancias i3
- Instancias i3en
- Instancias i4g
- Instancias i4i
- instancias i7i
- Instancias i7i
- instancias i8g
- Instancias im4gn
- Instancias is4gen
- Instancias m1
- Instancias m2
- Instancias m3
- Instancias m4
- Instancias m5
- Instancias m5a
- Instancias m5ad
- Instancias m5d
- Instancias m5dn
- Instancias m5n
- Instancias m5zn
- Instancias m6a
- Instancias m6g
- Instancias m6gd
- Instancias m6i
- Instancias m6id
- Instancias m6idn
- Instancias m6in
- Instancias m7a

- Instancias m7g
- Instancias m7gd
- Instancias m7i
- Instancias m7i-flex
- Instancias m8g
- instancias m8gd
- Instancias p2
- Instancias p3
- Instancias p4d
- Instancias p5
- Instancias r3
- Instancias r4
- Instancias r5
- Instancias r5a
- Instancias r5ad
- Instancias r5b
- Instancias r5d
- Instancias r5dn
- Instancias r5n
- Instancias r6a
- Instancias r6g
- Instancias r6gd
- Instancias r6i
- Instancias r6id
- Instancias r6idn
- Instancias r6in
- Instancias r7a
- Instancias r7g
- Instancias r7gd
- Instancias r7i

- Instancias r7iz
- Instancias r8g
- instancias r8gd
- Instancias x1
- Instancias x1e
- Instancias x2gd
- Instancias x2idn
- Instancias x2iedn
- Instancias x8g
- Instancias z1d

Instancias c1

c1.medium

Parámetro	Valor
YARN_RESOURCEMANAGER_HEAPSIZE	192
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	128
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

c1.xlarge

Parámetro	Valor
YARN_RESOURCEMANAGER_HEAPSIZE	768

Parámetro	Valor
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

Instancias c3

c3.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c3.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396

Parámetro	Valor
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

c3.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c3.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317

Parámetro	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

Instancias c4

c4.large

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1152
YARN_PROXYSERVER_HEAPSIZE	1152
YARN_NODEMANAGER_HEAPSIZE	1152
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1152
HADOOP_NAMENODE_HEAPSIZE	576
HADOOP_DATANODE_HEAPSIZE	384

c4.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124

Parámetro	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c4.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

c4.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c4.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

Instancias c5

c5.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5.9xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5.18xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c5a

c5a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c5ad

c5ad.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5ad.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5ad.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5ad.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5ad.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c5ad.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5ad.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c5d

c5d.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5d.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5d.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5d.9xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5d.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5d.18xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5d.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c5n

c5n.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2304
YARN_PROXYSERVER_HEAPSIZE	2304
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2304
HADOOP_NAMENODE_HEAPSIZE	1280
HADOOP_DATANODE_HEAPSIZE	665

c5n.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2519
YARN_PROXYSERVER_HEAPSIZE	2519
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2519
HADOOP_NAMENODE_HEAPSIZE	2355
HADOOP_DATANODE_HEAPSIZE	880

c5n.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2949
YARN_PROXYSERVER_HEAPSIZE	2949
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2949
HADOOP_NAMENODE_HEAPSIZE	4505
HADOOP_DATANODE_HEAPSIZE	1310

c5n.9xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5n.18xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c6a

c6a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6a.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c6a.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c6g

c6g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instancias c6gd

c6gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instancias c6gn

c6gn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instancias c6i

c6i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6i.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias C6id

c6id.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6id.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6id.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6id.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6id.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6id.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6id.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6id.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c6in

c6in.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6in.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6in.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6in.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6in.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6in.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6in.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6in.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c7a

c7a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7a.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c7a.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c7g

c7g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instancias c7gd

c7gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instancias C7gn

c7gn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instancias c7i

c7i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7i.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instancias c7i-flex

c7i-flex.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i-flex.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i-flex.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i-flex.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i-flex.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i-flex.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

Instancias c8g

c8g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8g.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8g.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

instancias c8gd

c8gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8gd.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8gd.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instancias d2

d2.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d2.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d2.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d2.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias d3

d3.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias d3en

d3en.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

d3en.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3en.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3en.6xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

d3en.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3en.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

instancias f2

f2.6xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

f2.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

f2.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instancias g3

g3.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g3.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g3.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias g3s

g3s.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

Instancias g4dn

g4dn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416

Parámetro	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

g4dn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

g4dn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

g4dn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

g4dn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

g4dn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instancias g5

g5.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

g5.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g5.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g5.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g5.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

g5.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g5.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g5.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias g6

g6.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

g6.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

g6.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

g6.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

g6.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5795
YARN_PROXYSERVER_HEAPSIZE	5795
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5795
HADOOP_NAMENODE_HEAPSIZE	18739
HADOOP_DATANODE_HEAPSIZE	4096

g6.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

g6.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9543
YARN_PROXYSERVER_HEAPSIZE	9543
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9543
HADOOP_NAMENODE_HEAPSIZE	37478
HADOOP_DATANODE_HEAPSIZE	4096

g6.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	16998
YARN_PROXYSERVER_HEAPSIZE	16998
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	16998
HADOOP_NAMENODE_HEAPSIZE	74752
HADOOP_DATANODE_HEAPSIZE	4096

Instancias g6e

g6e.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g6e.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g6e.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g6g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g6e.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g6e.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

g6e.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

g6e.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instancias gr6

gr6.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

gr6.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

Instancias h1

h1.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

h1.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

h1.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

h1.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instancias i2

i2.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i2.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i2.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i2.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias i3

i3.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i3.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i3.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i3.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i3.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias i3en

i3en.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

i3en.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

i3en.3xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

i3en.6xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

i3en.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

i3en.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instancias i4g

i4g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias i4i

i4i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i4i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i4i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i4i.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

instancias i7i

i7i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i7i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i7i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i7i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7i.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instancias i7i

i7ie.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7ie.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7ie.3xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

i7ie.6xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.18xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	13332
YARN_PROXYSERVER_HEAPSIZE	13332
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	13332
HADOOP_NAMENODE_HEAPSIZE	56422
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instancias i8g

i8g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i8g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i8g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i8g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i8g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i8g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i8g.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i8g.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instancias im4gn

im4gn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

im4gn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

im4gn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

im4gn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

im4gn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias is4gen

is4gen.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2557
YARN_PROXYSERVER_HEAPSIZE	2557
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2557
HADOOP_NAMENODE_HEAPSIZE	2547
HADOOP_DATANODE_HEAPSIZE	919

is4gen.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

is4gen.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

is4gen.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m1

m1.small

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	256
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	192
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

m1.medium

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	384
YARN_PROXYSERVER_HEAPSIZE	192
YARN_NODEMANAGER_HEAPSIZE	256

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	256
HADOOP_NAMENODE_HEAPSIZE	384
HADOOP_DATANODE_HEAPSIZE	192

m1.large

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	768
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

m1.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1024
YARN_PROXYSERVER_HEAPSIZE	512
YARN_NODEMANAGER_HEAPSIZE	768

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	2304
HADOOP_DATANODE_HEAPSIZE	384

Instancias m2

m2.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	3072
HADOOP_DATANODE_HEAPSIZE	384

m2.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	6144
HADOOP_DATANODE_HEAPSIZE	384

m2.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2048
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1536
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	12288
HADOOP_DATANODE_HEAPSIZE	384

Instancias m3

m3.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m3.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

Instancias m4

m4.large

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

m4.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m4.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m4.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m4.10xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5365
YARN_PROXYSERVER_HEAPSIZE	5365
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5365
HADOOP_NAMENODE_HEAPSIZE	16588
HADOOP_DATANODE_HEAPSIZE	3727

m4.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m5

m5.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m5a

m5a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m5ad

m5ad.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5ad.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5ad.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5ad.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5ad.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m5d

m5d.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5d.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5d.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5d.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5d.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5d.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5d.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m5dn

m5dn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5dn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5dn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5dn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5dn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m5n

m5n.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5n.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5n.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5n.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5n.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5n.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5n.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m5zn

m5zn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m5zn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5zn.3xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

m5zn.6xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

m5zn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m6a

m6a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6a.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m6a.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m6g

m6g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m6gd

m6gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m6i

m6i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6i.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m6id

m6id.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6id.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6id.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6id.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6id.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6id.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6id.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6id.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m6idn

m6idn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6idn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6idn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6idn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6idn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m6in

m6in.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6in.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6in.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6in.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6in.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6in.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6in.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6in.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m7a

m7a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7a.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m7a.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m7g

m7g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m7gd

m7gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m7i

m7i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m7i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7i.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m7i-flex

m7i-flex.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i-flex.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i-flex.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i-flex.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i-flex.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7i-flex.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias m8g

m8g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m8g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8g.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m8g.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

instancias m8gd

m8gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias p2

p2.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p2.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

p2.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias p3

p3.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p3.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

p3.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias p4d

p4d.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	24576
YARN_PROXYSERVER_HEAPSIZE	24576
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	24576
HADOOP_NAMENODE_HEAPSIZE	112640
HADOOP_DATANODE_HEAPSIZE	4096

Instancias p5

p5.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065

Parámetro	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r3

r3.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r3.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338

Parámetro	Valor
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r3.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r3.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r4

r4.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r4.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r4.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r4.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r4.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r5

r5.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r5a

r5a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r5ad

r5ad.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5ad.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5ad.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5ad.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12247
YARN_PROXYSERVER_HEAPSIZE	12247
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12247
HADOOP_NAMENODE_HEAPSIZE	50995
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r5b

r5b.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5b.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5b.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5b.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5b.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5b.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5b.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r5d

r5d.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5d.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5d.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5d.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5d.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5d.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5d.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r5dn

r5dn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5dn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5dn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5dn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r5n

r5n.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5n.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5n.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5n.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5n.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5n.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5n.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r6a

r6a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6a.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r6a.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r6g

r6g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r6gd

r6gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r6i

r6i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6i.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r6id

r6id.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6id.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6id.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6id.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6id.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6id.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6id.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6id.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r6idn

r6idn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6idn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6idn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6idn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r6in

r6in.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6in.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6in.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6in.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6in.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6in.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6in.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6in.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r7a

r7a.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7a.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7a.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7a.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7a.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7a.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7a.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7a.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r7a.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r7g

r7g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r7gd

r7gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r7i

r7i.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7i.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7i.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7i.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7i.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7i.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7i.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7i.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r7iz

r7iz.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7iz.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7iz.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7iz.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

Instancias r8g

r8g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8g.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r8g.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

instancias r8gd

r8gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

Instancias x1

x1.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

Instancias x1e

x1e.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4520
YARN_PROXYSERVER_HEAPSIZE	4520
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4520
HADOOP_NAMENODE_HEAPSIZE	12364
HADOOP_DATANODE_HEAPSIZE	2882

x1e.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6952
YARN_PROXYSERVER_HEAPSIZE	6952
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6952
HADOOP_NAMENODE_HEAPSIZE	24524
HADOOP_DATANODE_HEAPSIZE	4096

x1e.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	11816
YARN_PROXYSERVER_HEAPSIZE	11816
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	11816
HADOOP_NAMENODE_HEAPSIZE	48844
HADOOP_DATANODE_HEAPSIZE	4096

x1e.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1e.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

x1e.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	79912
YARN_PROXYSERVER_HEAPSIZE	79912
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	79912
HADOOP_NAMENODE_HEAPSIZE	389324
HADOOP_DATANODE_HEAPSIZE	4096

Instancias x2gd

x2gd.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x2gd.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2gd.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

Instancias x2idn

x2idn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

Instancias x2iedn

x2iedn.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2iedn.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.32xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	82042
YARN_PROXYSERVER_HEAPSIZE	82042
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	82042
HADOOP_NAMENODE_HEAPSIZE	399974
HADOOP_DATANODE_HEAPSIZE	4096

Instancias x8g

x8g.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x8g.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x8g.4xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x8g.8xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x8g.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x8g.16xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x8g.24xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x8g.48xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

Instancias z1d

z1d.xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

z1d.2xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

z1d.3xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

z1d.6xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

z1d.12xlarge

Parámetro	Valor
YARN_RESOURCEMANAG ER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048

Parámetro	Valor
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

Configuración de HDFS

La tabla siguiente describe los parámetros de Hadoop Distributed File System (HDFS) predeterminados y su configuración. Puede cambiar estos valores mediante la clasificación de configuración `hdfs-site`. Para obtener más información, consulte Configuración de aplicaciones.

Warning

1. Establecer `dfs.replication` en 1 en clústeres con menos de cuatro nodos puede conllevar la pérdida de datos del HDFS si un solo nodo deja de funcionar. Si su clúster tiene almacenamiento HDFS, le recomendamos que lo configure con al menos cuatro nodos principales para las cargas de trabajo de producción con el objetivo de evitar la pérdida de datos.
2. Amazon EMR no permitirá que los clústeres escalen los nodos principales por debajo de `dfs.replication`. Por ejemplo, si `dfs.replication = 2`, el número mínimo de nodos principales es 2.
3. Cuando utiliza el escalado administrado, emplea el escalado automático o decide cambiar el tamaño del clúster manualmente, se recomienda que establezca `dfs.replication` en 2 o más.

Parámetro	Definición	Valor predeterminado
`dfs.block.size`	El tamaño de bloques de HDFS. Cuando se opera con datos almacenados en HDFS, el tamaño dividido es por lo general el tamaño de un bloque de HDFS. Los números	134 217 728 (128 MB)

Parámetro	Definición	Valor predeterminado
	más grandes ofrecen menos granularidad de tarea, pero también aplican menos sobrecarga al clúster `NameNode`.	
`dfs.replication`	El número de copias de cada bloque que almacenar por razones de durabilidad. Amazon EMR establece este valor en función del número de nodos principales con los que se aprovisiona el clúster. Ajuste el valor para que se adapte a sus necesidades. Para sobrescribir el valor predeterminado, utilice la clasificación `hdfs-site` .	1 para clústeres aprovisionados con menos de cuatro nodos principales 2 para clústeres aprovisionados con menos de diez nodos principales 3 para el resto de clústeres

Cifrado transparente en el HDFS en Amazon EMR

El cifrado transparente se implementa mediante el uso de zonas de cifrado de HDFS, que son rutas de HDFS definidas por usted. Cada zona de cifrado tiene su propia clave, que se almacena en el servidor de claves especificado mediante la clasificación de configuración `hdfs-site`.

A partir de la versión 4.8.0 de Amazon EMR, puede utilizar configuraciones de seguridad de Amazon EMR para definir configuraciones de cifrado de datos para los clústeres de manera más sencilla. Las configuraciones de seguridad ofrecen ajustes para habilitar la seguridad de los datos en tránsito y de los datos en reposo en volúmenes de almacenamiento de Amazon Elastic Block Store (Amazon EBS) y datos de EMRFS en Amazon S3. Para obtener más información, consulte Cifrado de datos en tránsito y en reposo en la Guía de administración de Amazon EMR.

Amazon EMR usa el KMS de Hadoop de forma predeterminada; sin embargo, puede usar otro KMS que implemente la operación de la API. KeyProvider Cada archivo de una zona de cifrado HDFS tiene su propia clave de cifrado de datos exclusiva, que se cifra mediante la clave de zona de cifrado.

Los datos de HDFS se cifran end-to-end (en reposo y en tránsito) cuando se escriben en una zona de cifrado, ya que las actividades de cifrado y descifrado solo se llevan a cabo en el cliente.

No puede mover archivos entre zonas de cifrado o desde una zona de cifrado a las rutas no cifradas.

El cliente HDFS NameNode y el cliente HDFS interactúan con el KMS de Hadoop (o un KMS alternativo que haya configurado) a través de la operación de la API. KeyProvider El KMS es responsable de almacenar las claves de cifrado en el almacén de claves de copia de seguridad. Además, Amazon EMR incluye la política de seguridad ilimitada de JCE, por lo que puede crear claves con la longitud deseada.

Para obtener más información, consulte Cifrado transparente en el HDFS en la documentación de Hadoop.

> **Note**
>
> En Amazon EMR, KMS sobre HTTPS no está habilitado de forma predeterminada con Hadoop KMS. Para obtener más información sobre cómo habilitar KMS sobre HTTPS, consulte la documentación de Hadoop KMS.

Configuración del cifrado transparente de HDFS

Puede configurar el cifrado transparente en Amazon EMR mediante la creación de claves y la adición de zonas de cifrado. Puede hacer esto de varias formas:

- Uso de la operación del API de configuración de Amazon EMR al crear un clúster
- Mediante un paso de JAR de Hadoop con command-runner.jar
- Registrando en el nodo principal del clúster de Hadoop y utilizando los clientes de línea de comando `hadoop key` y `hdfs crypto`
- Uso del REST APIs para Hadoop KMS y HDFS

Para obtener más información sobre el REST APIs, consulte la documentación correspondiente a Hadoop KMS y HDFS.

Para crear zonas de cifrado y sus claves en la creación del clúster mediante la interfaz de línea de comandos (CLI)

La clasificación `hdfs-encryption-zones` en la operación de la API de configuración le permite especificar un nombre de clave y una zona de cifrado al crear un clúster. Amazon EMR crea esta clave en Hadoop KMS en su clúster y configura la zona de cifrado.

- Cree un clúster con el siguiente comando.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

`myConfig.json`:

```
[
 {
  "Classification": "hdfs-encryption-zones",
  "Properties": {
   "/myHDFSPath1": "path1_key",
   "/myHDFSPath2": "path2_key"
  }
 }
]
```

Para crear zonas de cifrado y sus claves manualmente en el nodo principal

1. Lance su clúster con una versión de Amazon EMR posterior a la 4.1.0.
2. Conéctese al nodo maestro del clúster con SSH.
3. Cree una clave en Hadoop KMS.

```
$ hadoop key create path2_key
path2_key has been successfully created with options Options{cipher='AES/CTR/
NoPadding', bitLength=256, description='null', attributes=null}.
KMSClientProvider[http://ip-x-x-x-x.ec2.internal:16000/kms/v1/] has been updated.
```

> **Important**
>
> Hadoop KMS requiere que sus nombres de clave estén en minúsculas. Si utiliza una clave que tiene mayúsculas entonces el clúster provocará un error durante el lanzamiento.

4. Cree la ruta de la zona de cifrado en HDFS.

```
$ hadoop fs -mkdir /myHDFSPath2
```

5. Haga de la ruta de HDFS una zona de cifrado con la clave que ha creado.

```
$ hdfs crypto -createZone -keyName path2_key -path /myHDFSPath2
Added encryption zone /myHDFSPath2
```

Para crear zonas de cifrado y sus claves manualmente mediante el AWS CLI

- Añada pasos para crear manualmente las zonas de cifrado y las claves de KMS con el siguiente comando.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Create
 First Hadoop KMS Key",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[/
bin/bash,-c,"\"hadoop key create path1_key\""] \
Type=CUSTOM_JAR,Name="Create First Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create First Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path1_key -path /myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop KMS Key",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop key create
 path2_key\""] \
```

```
Type=CUSTOM_JAR,Name="Create Second Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath2\""] \
Type=CUSTOM_JAR,Name="Create Second Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path2_key -path /myHDFSPath2\""]
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Consideraciones para cifrado transparente de HDFS

Una práctica recomendada consiste en crear una zona de cifrado para cada aplicación en la que pueda escribir archivos. Además, puede cifrar todos los HDFS mediante la hdfs-encryption-zones clasificación de la API de configuración y especificar la ruta raíz (/) como zona de cifrado.

Hadoop Key Management Server

Hadoop KMS es un servidor de administración de claves que ofrece la posibilidad de implementar servicios criptográficos para clústeres de Hadoop y puede funcionar como proveedor de claves para Cifrado transparente en el HDFS en Amazon EMR. Hadoop KMS en Amazon EMR se instala y habilita de forma predeterminada al seleccionar la aplicación de Hadoop al lanzar un clúster de EMR. El Hadoop KMS no almacena las claves en sí, salvo en el caso de almacenamiento temporal en caché. Hadoop KMS actúa como un proxy entre el proveedor de claves y el elemento de confianza cliente para un almacén de claves de respaldo: no es un almacén de claves. El almacén de claves predeterminado que se crea para Hadoop KMS es la extensión de criptografía de Java (JCEKS). KeyStore La política de seguridad ilimitada de JCE también se incluye, por lo que puede crear claves con la longitud deseada. Hadoop KMS también admite una gama de aplicaciones ACLs que controlan el acceso a las claves y las operaciones clave de forma independiente de otras aplicaciones cliente, como HDFS. La longitud de clave predeterminada en Amazon EMR es de 256 bits.

Para configurar Hadoop KMS, utilice la clasificación para cambiar la hadoop-kms-site configuración. Para configurarlo ACLs, utilice la clasificación kms-acls.

Para obtener más información, consulte la documentación de Hadoop KMS. Hadoop KMS se utiliza en el cifrado transparente de Hadoop HDFS. Para obtener más información sobre el cifrado transparente de HDFS, consulte el tema Cifrado transparente de HDFS en la documentación de Apache Hadoop.

> **Note**
>
> En Amazon EMR, KMS sobre HTTPS no está habilitado de forma predeterminada con Hadoop KMS. Para obtener información sobre cómo habilitar KMS sobre HTTPS, consulte la documentación de Hadoop KMS.

> **Important**
>
> Hadoop KMS requiere que sus nombres de clave estén en minúsculas. Si utiliza una clave que tiene mayúsculas entonces el clúster provocará un error durante el lanzamiento.

Configuración de Hadoop KMS en Amazon EMR

En la versión de lanzamiento de Amazon EMR 4.6.0 o posterior, `kms-http-port` es 9700 y `kms-admin-port` es 9701.

Puede configurar Hadoop KMS en el momento de la creación del clúster mediante la API de configuración para versiones de Amazon EMR. A continuación se enumeran las clasificaciones de objetos de configuración disponibles para Hadoop KMS:

Clasificaciones de configuración de Hadoop KMS

Clasificación	Filename
hadoop-kms-site	`kms-site.xml`
hadoop-kms-acls	`kms-acls.xml`
hadoop-kms-env	`kms-env.sh`
hadoop-kms-log4j	`kms-log4j.properties`

Para configurar Hadoop KMS ACLs mediante la CLI

- Cree un clúster con Hadoop KMS con el siguiente ACLs comando:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-acls",
  "Properties": {
   "hadoop.kms.blacklist.CREATE": "hdfs,foo,myBannedUser",
   "hadoop.kms.acl.ROLLOVER": "myAllowedUser"
  }
 }
]
```

Para deshabilitar la caché de Hadoop KMS a través de la interfaz de línea de comandos (CLI)

- Cree un clúster con Hadoop KMS `hadoop.kms.cache.enable` definido en `false` utilizando el siguiente comando:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-site",
  "Properties": {
   "hadoop.kms.cache.enable": "false"
  }
 }
]
```

Para configurar las variables de entorno en el script **`kms-env.sh`** mediante la interfaz de línea de comandos (CLI)

- Cambie la configuración a `kms-env.sh` través de la configuración `hadoop-kms-env`. Cree un clúster con Hadoop KMS utilizando el siguiente comando:

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-env",
  "Properties": {},
  "Configurations": [
   {
    "Classification": "export",
    "Properties": {
     "JAVA_LIBRARY_PATH": "/path/to/files",
     "KMS_SSL_KEYSTORE_FILE": "/non/Default/Path/.keystore",
     "KMS_SSL_KEYSTORE_PASS": "myPass"
    },
    "Configurations": []
   }
  ]
 }
]
```

Para obtener información acerca de la configuración de Hadoop KMS, consulte la documentación de Hadoop KMS.

Cifrado transparente de HDFS en clústeres de EMR con varios nodos maestros

Apache Ranger KMS se utiliza en un clúster de Amazon EMR con varios nodos principales para el cifrado transparente en el HDFS.

Apache Ranger KMS almacena su clave raíz y las claves de la zona de cifrado (EZ) en su Amazon RDS para un clúster de Amazon EMR con varios nodos principales. Para habilitar el cifrado transparente en el HDFS en un clúster de Amazon EMR con varios nodos principales, debe proporcionar las siguientes configuraciones.

- Amazon RDS o su propia URL de conexión de servidor MySQL para almacenar la clave principal de Ranger KMS y la clave de EZ
- Nombre de usuario y contraseña de MySQL
- Contraseña de la clave principal de Ranger KMS

- Archivo PEM de la entidad de certificación (CA) para la conexión SSL al servidor MySQL. Puede descargar el paquete de certificados para usted en Descargar paquetes Región de AWS de certificados para Amazon RDS.

Puede proporcionar estas configuraciones mediante las clasificaciones `ranger-kms-dbks-site` y `ranger-kms-db-ca`, tal y como se muestra en el siguiente ejemplo.

```
[
 {
  "Classification": "ranger-kms-dbks-site",
  "Properties": {
   "ranger.ks.jpa.jdbc.url": "jdbc:log4jdbc:mysql://mysql-host-url.xx-
xxx-1.xxx.amazonaws.com:3306/rangerkms",
   "ranger.ks.jpa.jdbc.user": "mysql-user-name",
   "ranger.ks.jpa.jdbc.password": "mysql-password",
   "ranger.db.encrypt.key.password": "password-for-encrypting-a-master-key"
  }
 },
 {
  "Classification": "ranger-kms-db-ca",
  "Properties": {
   "ranger.kms.trust.ca.file.s3.url": "<S3-path-of-downloaded-pem-file>"
  }
 }
]
```

A continuación se muestran las clasificaciones de objetos de configuración para Apache Ranger KMS.

Clasificaciones de configuración de Hadoop KMS

Clasificación	Descripción
ranger-kms-dbks-site	Cambie los valores del archivo dbks-site.xml de Ranger KMS.
ranger-kms-site	Cambie los valores del ranger-kms-site archivo.xml de Ranger KMS.
ranger-kms-env	Cambie los valores del entorno Ranger KMS.

Clasificación	Descripción
ranger-kms-log4j	Cambie los valores del archivo kms-log4j .properties de Ranger KMS.
ranger-kms-db-ca	Cambie los valores del archivo CA en S3 para la conexión SSL de MySQL con Ranger KMS.

Consideraciones

- Se recomienda encarecidamente cifrar la instancia de Amazon RDS para mejorar la seguridad. Para obtener más información, consulte Información general del cifrado de los recursos de Amazon RDS.
- Se recomienda encarecidamente utilizar una base de datos MySQL independiente en cada clúster de Amazon EMR con nodos principales para aumentar la seguridad.
- Para configurar el cifrado transparente en el HDFS en un clúster de Amazon EMR con varios nodos principales, debe especificar la clasificación `hdfs-encryption-zones` al crear el clúster. De lo contrario, Ranger KMS no se configurará ni iniciará. La reconfiguración de la clasificación `hdfs-encryption-zones` o cualquiera de las clasificaciones de configuración de Hadoop KMS en un clúster en ejecución no se admite en clúster de Amazon EMR con varios nodos principales.
- El paquete de certificados PEM que descarga de Descargar paquetes de certificados para Amazon RDS agrupa varios certificados en un solo archivo. Amazon EMR 7.3.0 y las versiones posteriores admiten la importación de varios certificados del archivo PEM con la configuración `ranger.kms.trust.ca.file.s3.url`.

Crear o ejecutar una aplicación de Hadoop

Temas

- Creación de binarios con Amazon EMR
- Procesamiento de datos con streaming
- Procesamiento de datos con un JAR personalizado

Creación de binarios con Amazon EMR

Puede utilizar Amazon EMR como un entorno de compilación para compilar programas para su uso en el clúster. Los programas que utilice con Amazon EMR deben compilarse en un sistema que ejecute la misma versión de Linux que utiliza Amazon EMR. Para una versión de 32 bits, debería haber compilado en una máquina de 32 bits o con las opciones de compilación cruzada de 32 bits activadas. Para una versión de 64 bits, tiene que haber compilado en una máquina de 64 bits o con las opciones de compilación cruzada de 64 bits activadas. Para obtener más información sobre las versiones de EC2 instancias, consulte Planificar y configurar EC2 instancias en la Guía de administración de Amazon EMR. Entre los lenguajes de programación admitidos se encuentran C++, Python y C#.

En la siguiente tabla se describen los pasos necesarios para crear y probar su aplicación mediante Amazon EMR.

Proceso de creación de un módulo

1	Conecte al nodo principal en su clúster.
2	Copie los archivos de origen al nodo principal.
3	Cree binarios con las optimizaciones necesarias.
4	Copie binarios desde el nodo maestro a Amazon S3.

Los detalles de cada uno de estos pasos se tratan en las siguientes secciones.

Para conectar al nodo principal del clúster

- Siga las instrucciones de Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.

Para copiar archivos de origen al nodo principal

1. Ponga sus archivos de origen en un bucket de Amazon S3. Para obtener información acerca de cómo crear buckets y cómo transferir datos a Amazon S3, consulte la Guía del usuario de Amazon Simple Storage Service.
2. Cree una carpeta en su clúster de Hadoop para los archivos de origen introduciendo un comando similar al siguiente:

```
mkdir SourceFiles
```

3. Copie sus archivos de origen de Amazon S3 al nodo maestro escribiendo un comando similar al siguiente:

```
hadoop fs -get s3://amzn-s3-demo-bucket/SourceFiles SourceFiles
```

Cree binarios con las optimizaciones necesarias

La forma de crear sus binarios depende de muchos factores. Siga las instrucciones de sus herramientas de compilación específicas para establecer y configurar su entorno. Puede utilizar comandos de especificaciones del sistema de Hadoop para obtener información de clústeres para determinar cómo instalar el entorno de compilación.

Para identificar las especificaciones del sistema

- Utilice los siguientes comandos para comprobar la arquitectura que está utilizando para crear sus binarios.

  a. Para ver la versión de Debian, escriba el siguiente comando:

  ```
  master$ cat /etc/issue
  ```

  El resultado tiene un aspecto similar al siguiente.

  ```
  Debian GNU/Linux 5.0
  ```

  b. Para ver el nombre de DNS pública y el tamaño de procesador, escriba el siguiente comando:

  ```
  master$ uname -a
  ```

  El resultado tiene un aspecto similar al siguiente.

  ```
  Linux domU-12-31-39-17-29-39.compute-1.internal 2.6.21.7-2.fc8xen #1 SMP Fri
   Feb 15 12:34:28 EST 2008 x86_64 GNU/Linux
  ```

  c. Para ver la velocidad del procesador, escriba el siguiente comando:

```
master$ cat /proc/cpuinfo
```

El resultado tiene un aspecto similar al siguiente.

```
processor : 0
vendor_id : GenuineIntel
model name : Intel(R) Xeon(R) CPU E5430 @ 2.66GHz
flags : fpu tsc msr pae mce cx8 apic mca cmov pat pse36 clflush dts acpi mmx
 fxsr sse sse2 ss ht tm syscall nx lm constant_tsc pni monitor ds_cpl vmx est
 tm2 ssse3 cx16 xtpr cda lahf_lm
...
```

Una vez que se hayan creado binarios, podrá copiar los archivos en Amazon S3.

Para copiar binarios desde el nodo maestro a Amazon S3

- Escriba el siguiente comando para copiar los binarios a su bucket de Amazon S3:

```
hadoop fs -put BinaryFiles s3://amzn-s3-demo-bucket/BinaryDestination
```

Procesamiento de datos con streaming

El streaming de Hadoop es una utilidad que viene con Hadoop y que le permite desarrollar MapReduce ejecutables en lenguajes que no sean Java. El streaming se implementa en forma de un archivo JAR, por lo que puede ejecutarlo desde la API de Amazon EMR o la línea de comando como un archivo JAR estándar.

En esta sección se describe cómo utilizar el streaming con Amazon EMR.

> **Note**
>
> El streaming de Apache Hadoop es una herramienta independiente. Por tanto, todas las funciones y parámetros no se describen aquí. Para más información sobre el streaming de Hadoop, vaya a http://hadoop.apache. org/docs/stable/hadoop-streaming/HadoopStreaming.html.

Uso de la utilidad de streaming de Hadoop

Esta sección describe cómo utilizar la utilidad de streaming de Hadoop.

Proceso de Hadoop

1	Escriba su ejecutable de mapeador y reductor en el lenguaje de programación que desee. Siga las instrucciones de la documentación de Hadoop para escribir sus ejecutables de streaming. Los programas deben leer su entrada a partir de los datos de entrada y salida estándar a través de la salida estándar. De forma predeterminada, cada línea de entrada/salida representa un registro y la primera pestaña de cada línea se utiliza como separador entre la clave y el valor.
2	Pruebe sus ejecutables localmente y cárguelos en Amazon S3.
3	Utilice la interfaz de línea de comandos de Amazon EMR o la consola de Amazon EMR para ejecutar su aplicación.

Cada script de mapeador se lanza como proceso independiente en el clúster. Cada ejecutable del reductor convierte la salida del ejecutable del mapeador en la salida de datos del flujo de trabajo.

La mayoría de las aplicaciones de streaming requieren los parámetros `input`, `output`, `mapper` y `reducer`. En la siguiente tabla se describen estos y otros parámetros opcionales.

Parámetro	Descripción	Obligatorio
-input	Ubicación de los datos de entrada en Amazon S3. Tipo: cadena Valor predeterminado: None Restricción: URI. Si no se especifica ningún protocolo, utiliza el sistema de archivos predeterminado del clúster.	Sí
-output		Sí

Parámetro	Descripción	Obligatorio
	Ubicación en Amazon S3 donde Amazon EMR carga los datos procesados. Tipo: cadena Valor predeterminado: None Restricción: URI Valor predeterminado: si no se especifica una ubicación, Amazon EMR carga los datos en la ubicación especificada por `input`.	
-mapper	Nombre del ejecutable del mapeador. Tipo: cadena Valor predeterminado: None	Sí
-reducer	Nombre del ejecutable del reductor. Tipo: cadena Valor predeterminado: None	Sí
-cacheFile	Una ubicación de Amazon S3 que contiene archivos que Hadoop copia en su directorio de trabajo local (principalmente para mejorar el rendimiento). Tipo: cadena Valor predeterminado: None Restricciones: [URI] # [nombre de symlink que crear en el directorio de trabajo]	No

Parámetro	Descripción	Obligatorio
-cacheArchive	Archivo JAR que extraer en el directorio de trabajo Tipo: cadena Valor predeterminado: None Restricciones: [URI] # [nombre de directorio de symlink que crear en el directorio de trabajo]	No
-combiner	Combina resultados Tipo: cadena Valor predeterminado: None Limitaciones: nombre de clase Java	No

El siguiente ejemplo de código es un ejecutable de mapeador escrito en Python. Este script forma parte de la aplicación de WordCount ejemplo.

```
#!/usr/bin/python
import sys

def main(argv):
  line = sys.stdin.readline()
  try:
    while line:
      line = line.rstrip()
      words = line.split()
      for word in words:
        print "LongValueSum:" + word + "\t" + "1"
      line = sys.stdin.readline()
  except "end of file":
    return None
if __name__ == "__main__":
  main(sys.argv)
```

Enviar un paso de streaming

En esta sección se describen los conceptos básicos de enviar un paso de streaming a un clúster. Una aplicación de streaming lee entradas desde la entrada estándar y, a continuación, ejecuta un script o ejecutable (denominado asignador) en contraste con cada entrada. El resultado de cada una de las entradas se guarda a nivel local, normalmente en una partición de Hadoop Distributed File System (HDFS). Después de que el mapeador procese todas las entradas, un segundo script o ejecutable (denominado reductor) procesa los resultados del mapeador. Los resultados del reductor se envían a la salida estándar. Puede encadenar conjuntamente una serie de pasos de streaming, donde la salida de un paso se convierte en la entrada de otro.

Se puede hacer referencia al mapeador y al reductor como un archivo o puede proporcionar una clase Java. Puede implementar el mapeador y reductor en cualquiera de los lenguajes admitidos, incluidos Ruby, Perl, Python, PHP o Bash.

Enviar un paso de streaming mediante la consola

En este ejemplo se describe cómo utilizar la consola de Amazon EMR para enviar un paso de streaming a un clúster en ejecución.

Para enviar un paso de streaming

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En la Cluster List (Lista de clústeres), seleccione el nombre del clúster.
3. Desplácese hasta la sección Steps (Pasos) y amplíela; a continuación, elija Add step (Añadir paso).
4. En el cuadro de diálogo Add Step (Añadir paso):
   - En Step type (Tipo de paso), elija Streaming program (Programa de streaming).
   - En Name (Nombre), acepte el nombre predeterminado, Streaming program (Programa de streaming), o escriba un nombre nuevo.
   - En Mapper (Mapeador), escriba o busque la ubicación de la clase de mapeador de Hadoop o del bucket de S3 donde reside el ejecutable del mapeador, por ejemplo, un programa de Python. El valor de la ruta debe tener el formato*BucketName*//. *path* *MapperExecutable*
   - En Reducer (Reductor), escriba o busque la ubicación de la clase de reductor de Hadoop o de un bucket de S3 donde resida el ejecutable del reductor, por ejemplo, un programa de Python. El valor de la ruta debe tener el formato*BucketName*/*path*/*MapperExecutable*. Amazon

EMR soporta la palabra clave especial agregación. Para obtener más información, vaya a la biblioteca Aggregate suministrada por Hadoop.

- En Input S3 location (Ubicación en S3 de la entrada), escriba o busque la ubicación de los datos de entrada.
- En Ubicación en S3 de la salida, escriba o busque el nombre del bucket de salida de Amazon S3.
- En Arguments (Argumentos), deje el campo en blanco.
- En Action on failure (Acción sobre el error), acepte la opción predeterminada, Continue (Continuar).

5. Elija Agregar. El paso aparece en la consola con el estado Pending (Pendiente).
6. El estado del paso cambia de Pending (Pendiente) a Running (En ejecución) y a Completed (Completado) a medida que se ejecuta. Para actualizar el estado, elija el icono Refresh (Actualizar) situado encima de la columna Actions (Acciones).

AWS CLI

Estos ejemplos muestran cómo utilizar el paso AWS CLI para crear un clúster y enviar un paso de streaming.

Para crear un clúster y enviar un paso de streaming mediante la AWS CLI

- Para crear un clúster y enviar un paso de streaming mediante la AWS CLI, escriba el siguiente comando y *myKey* sustitúyalo por el nombre del par de EC2 claves. Tenga en cuenta que su argumento para `--files` debe ser la ruta de Amazon S3 a la ubicación del script y los argumentos para `-mapper` y `-reducer` deben ser los nombres de los respectivos archivos de script.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 --
applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=STREAMING,Name="Streaming Program",ActionOnFailure=CONTINUE,Args=[--
files,pathtoscripts,-mapper,mapperscript,-reducer,reducerscript,aggregate,-
input,pathtoinputdata,-output,pathtooutputbucket]
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos secundarios. Todos los nodos utilizan el tipo de instancia especificado en el comando.

> **Note**
>
> Si no ha creado con anterioridad el rol de servicio de Amazon EMR predeterminado y el perfil de EC2 instancia, escriba aws `emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Procesamiento de datos con un JAR personalizado

Un JAR personalizado ejecuta un programa Java compilado que carga en Amazon S3. Debería compilar el programa en comparación con la versión de Hadoop que desee lanzar y envíe un paso `CUSTOM_JAR` a su clúster de Amazon EMR. Para más información sobre la compilación de un archivo JAR, consulte Creación de binarios con Amazon EMR.

Para obtener más información sobre cómo crear una MapReduce aplicación de Hadoop, consulte el MapReduce Tutorial en la documentación de Apache Hadoop.

Temas

- Enviar un paso JAR personalizado

Enviar un paso JAR personalizado

Un JAR personalizado ejecuta un programa Java compilado que carga en Amazon S3. Debería compilar el programa en comparación con la versión de Hadoop que desee lanzar y envíe un paso `CUSTOM_JAR` a su clúster de Amazon EMR. Para más información sobre la compilación de un archivo JAR, consulte Creación de binarios con Amazon EMR.

Para obtener más información sobre cómo crear una MapReduce aplicación de Hadoop, consulte el MapReduce Tutorial en la documentación de Apache Hadoop.

En esta sección se describen los conceptos básicos de enviar un paso JAR personalizado en Amazon EMR. Enviar un paso JAR personalizado le permite escribir un script para procesar los datos mediante el lenguaje de programación Java.

Enviar un paso JAR personalizado mediante la consola

En este ejemplo se describe cómo utilizar la consola de Amazon EMR para enviar un paso JAR personalizado a un clúster en ejecución.

Para enviar un paso JAR personalizado mediante la consola

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En la Cluster List (Lista de clústeres), seleccione el nombre del clúster.
3. Desplácese hasta la sección Steps (Pasos) y amplíela; a continuación, elija Add step (Añadir paso).
4. En el cuadro de diálogo Add Step (Añadir paso):
   - En Step type (Tipo de paso), elija Custom JAR (JAR personalizado).
   - En Name (Nombre), acepte el nombre predeterminado, Custom JAR (JAR personalizado), o escriba un nombre nuevo.
   - En JAR S3 location (Ubicación en S3 del JAR), escriba o busque la ubicación del archivo JAR. La ubicación del JAR puede ser una ruta hacia S3 o una clase java totalmente calificada en la classpath.
   - En Arguments (Argumentos), escriba los argumentos necesarios como cadenas separadas por espacios o deje el campo en blanco.
   - En Action on failure (Acción sobre el error), acepte la opción predeterminada, Continue (Continuar).
5. Elija Agregar. El paso aparece en la consola con el estado Pending (Pendiente).

6. El estado del paso cambia de Pending (Pendiente) a Running (En ejecución) y a Completed (Completado) a medida que se ejecuta. Para actualizar el estado, elija el icono Refresh (Actualizar) situado encima de la columna Actions (Acciones).

Inicio de un clúster y envío de un paso JAR personalizado mediante la AWS CLI

Para iniciar un clúster y enviar un paso JAR personalizado mediante la AWS CLI

Para iniciar un clúster y enviar un paso JAR personalizado mediante la AWS CLI, escriba el `create-cluster` subcomando con el `--steps` parámetro.

- Para iniciar un clúster y enviar un paso JAR personalizado, escriba el siguiente comando, *myKey* sustituya por el nombre del EC2 key pair y *amzn-s3-demo-bucket* sustituya por el nombre del bucket.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbuck
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se inicia un nodo principal único y el resto de las instancias se inician como nodos básicos. Todos los nodos utilizan el tipo de instancia que especifica en el comando.

> **Note**
>
> Si no ha creado con anterioridad el rol de servicio de Amazon EMR predeterminado y el perfil de EC2 instancia, escriba `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Dependencias de terceros

A veces puede ser necesario incluirlo en la MapReduce ruta de clases JARs para usarlo con el programa. Dispone de dos opciones para hacerlo:

- Incluir el paso `--libjars s3://`*`URI_to_JAR`* en las opciones de paso para el procedimiento en Inicio de un clúster y envío de un paso JAR personalizado mediante la AWS CLI.
- Lance el clúster con una configuración `mapreduce.application.classpath` modificada en `mapred-site.xml`. Utilice la clasificación de configuración `mapred-site`. Para crear el clúster con el paso mediante la AWS CLI, el aspecto sería el siguiente:

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--instance-type m5.xlarge --instance-count 2  --ec2-attributes KeyName=myKey \
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbucket
 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

`myConfig.json`:

```
[
    {
      "Classification": "mapred-site",
      "Properties": {
        "mapreduce.application.classpath": "path1,path2"
      }
    }
  ]
```

La lista de rutas separadas por comas se debe adjuntar al classpath para el JVM de cada tarea.

Leer objetos restaurados

Con Amazon EMR versión 7.2.0 y versiones posteriores, puede leer los objetos Glacier restaurados desde la ubicación S3 de la tabla con el protocolo `S3A`. Los motores de las versiones anteriores no distinguían entre los archivos de Glacier y los de Glacier Deep Archive, lo que significa que obtendría un `AmazonS3Exception` si intentara acceder a un archivo de Glacier en curso con `S3A`. Esta operación de lectura ignora los archivos de Glacier archivados si aún se están restaurando. Para habilitar este comportamiento, usa la configuración `fs.s3a.glacier.read.restored.objects`. Esta configuración puede tener tres valores:

- READ_ALL: este valor indica que Amazon EMR no debe tener en cuenta las clases de almacenamiento recuperadas de Amazon S3. Este es el comportamiento predeterminado.
- SKIP_ALL_GLACIER: este valor indica que Amazon EMR debe ignorar todos los objetos de S3 que estén etiquetados con la clase de almacenamiento Glacier y recuperar todos los demás objetos. Este es el comportamiento predeterminado para Amazon Athena con respecto a los objetos de Glacier.
- READ_RESTORED_GLACIER_OBJECTS: este valor indica que Amazon EMR debe comprobar el estado restaurado del objeto Glacier. Si Amazon EMR puede restaurar el objeto, puede leerlo como un objeto S3 normal. De lo contrario, Amazon EMR ignora el objeto de S3 Glacier.

Ejemplos

Spark

Para leer los objetos restaurados al utilizar Spark, utilice la siguiente configuración:

```
--conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

Si usa spark-sql, utilice la siguiente configuración en su lugar:

```
spark-sql --conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

Flink

Si usa Flink, puede establecer la configuración en el archivo `flink-conf.yaml`

```
fs.s3a.glacier.read.restored.objects: <value>
```

También puede configurar la clasificación `flink-conf`:

```
[
 {
  "Classification": "flink-conf",
  "Properties": {
  "fs.s3a.glacier.read.restored.objects":"<value>"
  }
 }
]
```

Hive

Si usa Hive, defina la configuración en el archivo `hive-site.xml`.

```
<property>
 <name>fs.s3a.glacier.read.restored.objects</name>
 <value><value></value>
</property>
```

También puede utilizar la CLI de Hive para configurar la propiedad `--hiveconf`:

```
hive --hiveconf fs.s3a.glacier.read.restored.objects=<value>
```

Consideraciones

Cuando lea objetos restaurados de S3 Glacier, tenga en cuenta las siguientes consideraciones:

- Solo puede leer los objetos restaurados si utiliza el esquema `S3A` o el `S3AFileSystem` para acceder a los datos.
- Al leer un objeto Glacier restaurado, Amazon EMR no restaura el objeto en sí. Para ello, debe utilizar el AWS CLI o el AWS SDK.

Activar el reconocimiento de acceso no uniforme a la memoria para los contenedores YARN

Con las versiones 6.x y posteriores de Amazon EMR, puede utilizar el acceso a memoria no uniforme (NUMA) para el multiprocesamiento de sus datos en clústeres. NUMA es un patrón de diseño de

memoria de computadora en el que el procesador puede acceder a su propia memoria local más rápido que a la memoria de otro procesador o la memoria compartida entre procesadores. Los contenedores YARN tienen un mejor rendimiento con NUMA porque pueden enlazarse a un nodo de NUMA específico que sirve para todas las asignaciones de memoria posteriores. Esto reduce la cantidad de veces que el clúster tiene que acceder a la memoria remota.

Puede activar la compatibilidad con NUMA para el contenedor YARN si la máquina del nodo de trabajo es un nodo de NUMA múltiple. Para confirmar si un nodo de trabajo es un nodo de NUMA único o NUMA múltiple, ejecute el siguiente comando.

```
lscpu | grep -i numa
NUMA node(s): 2
```

En general, las instancias que superan su tamaño 12 veces tienen dos nodos de NUMA. Esto no se aplica a las instancias de metal.

Para activar el conocimiento de NUMA sobre los contenedores YARN

1. Utilice la siguiente configuración de `yarn-site` en su clúster de Amazon EMR 6.x.

```
 [
 {
  "classification":"yarn-site",
   "properties":{
    "yarn.nodemanager.linux-container-executor.nonsecure-mode.local-user":"yarn",
    "yarn.nodemanager.linux-container-executor.group":"yarn",
    "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor",
    "yarn.nodemanager.numa-awareness.enabled":"true",
    "yarn.nodemanager.numa-awareness.numactl.cmd":"/usr/bin/numactl",
    "yarn.nodemanager.numa-awareness.read-topology":"true"
   },
  "configurations":[]
 }
 ]
```

2. Proporcione la siguiente acción de arranque en su clúster.

```
#!/bin/bash

sudo yum -y install numactl
```

```
echo 1 | sudo tee /proc/sys/kernel/numa_balancing

echo "banned.users=mapred,bin,hdfs" >> /etc/hadoop/conf/container-executor.cfg
rm -rf /var/log/hadoop-yarn/
sudo chown -R yarn:hadoop /var/log/hadoop-yarn/
sudo chmod 755 -R /var/log/hadoop-yarn/

sudo chmod 6050 /etc/hadoop/conf/container-executor.cfg

mkdir /mnt/yarn && sudo chmod 755 -R /mnt/yarn && sudo chown -R yarn:hadoop /mnt/
yarn
mkdir /mnt1/yarn && sudo chmod 755 -R /mnt1/yarn && sudo chown -R yarn:hadoop /
mnt1/yarn
mkdir /mnt2/yarn && sudo chmod 755 -R /mnt2/yarn && sudo chown -R yarn:hadoop /
mnt2/yarn
```

3. Todos los contenedores deben tener conocimiento de NUMA. Puede notificarlo a la máquina virtual de Java (JVM) de cada contenedor con una marca NUMA. Por ejemplo, para notificar a la JVM que utilice NUMA en un trabajo de MapReduce, agregue las siguientes propiedades en `mapred-site.xml`.

```
<property>
 <name>mapreduce.reduce.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
<property>
 <name>mapreduce.map.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
```

4. Para comprobar que haya activado NUMA, busque en cualquiera de los archivos de registro del NodeManager con el siguiente comando.

```
grep "NUMA resources allocation is enabled," *
```

Para comprobar si NodeManager ha asignado recursos de NUMA nodo a un contenedor, busque en el NodeManager registro con el siguiente comando y sustitúyalo por *<container_id>* su propio ID de contenedor.

```
grep "NUMA node" | grep <container_id>
```

Embalaje de contenedores YARN

A partir de la versión 7.9.0 de Amazon EMR, la política de embalaje de contenedores ya está disponible para el programador de capacidad de YARN, que se basa en la política de ubicación de varios nodos de YARN. Si bien la función está deshabilitada de forma predeterminada, cuando está activada, YARN prioriza llenar un solo nodo con contenedores antes de expandirlo a otros nodos del clúster, respetando un umbral de empaquetado predefinido definido por la configuración. `yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage`

La política de empaquetado de contenedores ofrece varias ventajas en comparación con la estrategia de asignación uniforme de contenedores predeterminada:

- Reduce la fragmentación de los recursos del clúster.
- Potencialmente, acelera las operaciones de reducción de escala de los clústeres al lanzar contenedores en un número limitado de nodos cuando hay recursos disponibles en esos nodos y, por lo tanto, dejar otros nodos inactivos, que luego se pueden reducir, lo que se traduce en un mayor ahorro de costes al escalar dinámicamente un clúster.

Habilite la función

Para habilitar la función de embalaje de contenedores en Amazon EMR, puede añadir la siguiente clasificación de sitios de YARN:

```
[
 {
  "Classification": "yarn-site",
  "Properties": {
  "yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage":
 "integer value from 1-100"
  }
 }
]
```

Consideraciones

- La función está disponible exclusivamente para el programador de capacidad YARN.

- Al habilitar la función, se activa automáticamente la estrategia de programación de ubicación de varios nodos de YARN.
- Puede haber una posible degradación del rendimiento debido a la utilización concentrada de los recursos en un número limitado de nodos.
- Con esta función, las políticas de autoescalado personalizadas demuestran mejores operaciones de escalado descendente, en comparación con la política de escalado gestionado.

Historial de versiones de Hadoop

En la siguiente tabla, se muestra la versión de Hadoop incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Hadoop

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-7.9.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.8.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
		de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.7.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.6.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-7.5.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.4.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-7.3.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.2.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.36.2	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.1.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-7.0.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.15.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.14.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.13.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.12.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.11.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.11.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.10.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.10.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.9.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.9.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.8.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.8.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.7.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.36.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.36.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.6.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.35.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.5.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.4.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.3.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.3.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.2.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.2.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.1.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.1.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-6.0.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.0.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.34.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.33.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.33.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.32.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.32.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.31.1	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.31.0	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.30.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.29.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.28.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.28.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.27.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.27.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.26.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.25.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.24.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.24.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.23.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.23.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.22.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.21.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.21.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.21.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.20.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.20.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.19.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.19.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.18.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.18.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.2	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.17.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.16.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.16.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.15.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.15.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.14.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.14.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.14.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.13.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.13.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.3	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.12.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.12.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.11.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.11.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.10.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.10.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.9.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.8.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.7.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.7.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.6.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.6.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.5.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.5.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.5.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.3.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.2.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.2.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.0.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-5.0.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.6	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.9.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.8.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.4	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.3	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.7.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.6.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.6.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.5.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.4.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.3.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.2.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Etiqueta de versión de Amazon EMR	Versión de Hadoop	Componentes instalados con Hadoop
emr-4.1.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.0.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Notas de la versión de Hadoop por versión

Temas

- Amazon EMR 7.9.0: notas de la versión de Hadoop
- Amazon EMR 7.8.0: notas de la versión de Hadoop
- Amazon EMR 7.7.0: notas de la versión de Hadoop
- Amazon EMR 7.6.0: notas de la versión de Hadoop
- Amazon EMR 7.5.0: notas de la versión de Hadoop
- Amazon EMR 7.4.0: notas de la versión de Hadoop
- Amazon EMR 7.3.0: notas de la versión de Hadoop
- Amazon EMR 7.2.0: notas de la versión de Hadoop

- Amazon EMR 6.6.0: notas de la versión de Hadoop

Amazon EMR 7.9.0: notas de la versión de Hadoop

Amazon EMR 7.9.0: cambios en Hadoop

Tipo	Descripción
Nueva característica	Asignación automática de la configuración: la configuración del sistema de archivos EMRFS se aplica automáticamente a la configuración del sistema de archivos S3A correspondiente, lo que permite una migración sin problemas de EMRFS a S3A.
Nueva característica	Empaquetado en contenedores YARN: política de programación para una reducción de escala agresiva
Portabilidad con versiones anteriores	YARN-11752: Programador global: mejora el tiempo de asignación de contenedores
Corrección de errores	S3A: Se corrige el problema de que el proveedor devuelva las credenciales IAMCredentials
Corrección de errores	Se corrigió el valor `yarn.log.server.url` de configuración para los clústeres con el cifrado en tránsito activado

Amazon EMR 7.8.0: notas de la versión de Hadoop

Amazon EMR 7.8.0: cambios en Hadoop

Tipo	Descripción
Upgrade	Hadoop se ha actualizado a la versión 3.4.1.

Amazon EMR 7.7.0: notas de la versión de Hadoop

Amazon EMR 7.7.0: cambios en Hadoop

Tipo	Descripción
Nueva característica	Optimice las llamadas de S3A con la lista de prefijos GlobStatus de S3
Portabilidad con versiones anteriores	YARN-7327: Habilita la programación asíncrona de forma predeterminada para el programador de capacidad
Portabilidad con versiones anteriores	YARN-10058: Gestiona la excepción no detectada en los subprocesos de programación asíncrona para evitar que el programador se bloquee
Portabilidad con versiones anteriores	YARN-11732: Se corrige un posible NPE al llamar a #reservedContainer para Scheduler Node CapacityScheduler
Portabilidad con versiones anteriores	YARN-11560: Se corrige el error de NPE cuando se activaban varios nodos con una programación asíncrona

Tipo	Descripción
Portabilidad con versiones anteriores	YARN-11191: Se corrige un posible GlobalSch eduler punto muerto en RefreshQueues
Portabilidad con versiones anteriores	YARN-11041: Reemplaza todas las apariciones de QueuePath por la nueva clase QueuePath
Portabilidad con versiones anteriores	YARN-11660: Se corrige la regresión del rendimiento para SingleConstraintAp pPlacementAllocator
Portabilidad con versiones anteriores	HADOOP-19116: actualización a la versión 3.8.4 del cliente zookeeper debido al CVE-2024-23944.
Portabilidad con versiones anteriores	HADOOP-19115: actualización a la nimbus-jose-jwt versión 9.37.2 debido al CVE-2023-52428.
Portabilidad con versiones anteriores	HADOOP-19024: Usa el bouncycastle jdk18 1.77
Portabilidad con versiones anteriores	HADOOP-19123: actualización a commons-c onfiguration2 2.10.1 debido al CVE
Portabilidad con versiones anteriores	HADOOP-19114: Actualización a commons-c ompress 1.26.1 debido a CVEs
Portabilidad con versiones anteriores	HADOOP-19237: Actualización a dnsjava 3.6.1 debido a CVEs
Nueva característica	Añada la auditoría de solicitudes de S3 a S3A

Tipo	Descripción
Portabilidad con versiones anteriores	HADOOP-17609: Hacer que la compatibilidad con el código nativo de SM4 OpenSSL sea opcional
Portabilidad con versiones anteriores	HADOOP-18583: hadoop checknative no puede cargar openssl 3.x
Nueva característica	Se agregó soporte para los mapeos de roles de S3A

Amazon EMR 7.7.0: Características de Hadoop

- La programación asíncrona de contenedores se ha convertido en la estrategia de programación predeterminada para el programador de capacidad, diseñada para optimizar la velocidad de asignación de contenedores.
- El sistema de archivos S3A introduce una optimización para las llamadas de estado global utilizando la lista de prefijos S3 para acelerar las operaciones de la lista. De forma predeterminada, esta función está deshabilitada y se puede activar `fs.s3a.prefix.listing.in.glob.status.enabled=true` configurándola en el archivo core-site.xml. Cuando está habilitada, la optimización permite filtrar desde el lado del servidor las llamadas a globstatus`fs.globstatus("s3://bucket/a*")`, lo que mejora el rendimiento de la lista al enumerar solo los objetos que comienzan por. `"a"`
- Agregue la auditoría de solicitudes de S3 a S3A. Cuando esté habilitada, la información del fileSystemOwner objeto se utilizará para rellenar la cadena UserAgent con los campos de usuario y grupo de usuarios que realizan las solicitudes de S3.
- El S3A añade compatibilidad con las asignaciones de funciones, lo que ayuda a determinar qué función de IAM utilizar en función de los usuarios, los grupos o los prefijos de S3.

Amazon EMR 7.6.0: notas de la versión de Hadoop

Amazon EMR 7.6.0: cambios en Hadoop

Tipo	Descripción
Nueva característica	Marco de la venta de credenciales S3A
Nueva característica	S3A Magic Committer Support para Hive
Portabilidad con versiones anteriores	HADOOP-19309: S3A: la CopyFromLocalFile operación falla cuando el archivo fuente no contiene el esquema de archivos
Portabilidad con versiones anteriores	HADOOP-19050 S3A: Compatible con las concesiones de acceso S3
Portabilidad con versiones anteriores	HADOOP-18708 S3A: Support S3 Client Side Encryption (CSE)

Amazon EMR 7.6.0: Características de Hadoop

- Los ajustes de configuración predeterminados para YARN se modificaron `yarn.scheduler.increment-allocation-mb` y `yarn.scheduler.minimum-allocation-mb` pasaron de 32 a 1, lo que deshabilita de forma efectiva la normalización de los recursos del contenedor en YARN.
- El sistema de archivos S3A de Hadoop se ha mejorado para admitir la venta de credenciales a nivel de solicitud. Con esta función, cada solicitud de Amazon S3 realizada utilizará un`AwsCredentialsProvider`. Esto garantiza una mayor flexibilidad y seguridad al proporcionar credenciales de forma dinámica para cada solicitud.
- El sistema de archivos S3A de Hadoop admite el cifrado S3 del lado del cliente.

Amazon EMR 7.5.0: notas de la versión de Hadoop

Amazon EMR 7.5.0: cambios en Hadoop

Tipo	Descripción
Corrección de errores	Se ha comentado fs.file.impl como valor vacío.
Portabilidad con versiones anteriores	HADOOP-19286: Support S3A entre regiones cuando S3 está configurado region/endpoint
Mejora	Configuración automática de la región S3 para el conector S3A en EMR- EC2
Mejora	Reduzca la cantidad de HeadObject llamadas en el S3A

Con el lanzamiento de Amazon EMR 7.5, el conector S3A de Spark demuestra un rendimiento de lectura comparable al de EMRFS, como lo demuestran los puntos de referencia que utilizan un conjunto de datos de parquet TPC-DS de 3 TB.

Amazon EMR 7.5.0: Características de Hadoop

- `fs.s3a.endpoint.region`La configuración de la región S3 se establece automáticamente en la región en la que se lanza el clúster de EMR con el conector S3A para la implementación de EMR. EC2
- El acceso entre regiones de compartimentos de Amazon S3 está habilitado de forma predeterminada para el conector S3A. Se puede modificar mediante la configuración. `fs.s3a.cross.region.access.enabled=true or false`

Amazon EMR 7.4.0: notas de la versión de Hadoop

Amazon EMR 7.4.0: cambios en Hadoop

Tipo	Descripción
Upgrade	La versión de Hadoop se ha actualizado a la 3.4.0; consulte las notas de la versión de OSS.
Corrección de errores	Corrija las métricas negativas de Yarn pendientes y asignadas para FairScheduler
Corrección de errores	YARN-11702: Soluciona el problema de Yarn por encima de la asignación de contenedores
Corrección de errores	Mejora el manejo de las condiciones de carrera al reducir el tamaño de los nodos
Mejora	HADOOP-18679: Añade una API para eliminar archivos bulk/paged
Mejora	HADOOP-19203: API de WrappeDio para mostrarla como no verificada BulkDelete IOEs IOExceptions
Mejora	HADOOP-19205: S3A: más lento que con el SDK de la versión 1 initialization/close
Mejora	HADOOP-19161: S3A: opción fs.s3a.performance.flags para incluir una lista de indicadores de rendimiento
Mejora	

Tipo	Descripción
	HADOOP-19072: S3A: amplía las optimizaciones de las tiendas con fs.s3a.performance.flags para mkdir

Amazon EMR 7.4.0: Características de Hadoop

Consulte la siguiente lista para ver las nuevas funciones de Hadoop en Amazon EMR 7.4.0.

- Los valores de configuración predeterminados se han ajustado para lograr un rendimiento óptimo:
  - mapreduce.input.fileinputformat.list-status.num-threads=10: esto es superior a 1.
  - fs.s3a.block.size=64 m: es un aumento en comparación con 32 millones.
  - fs.S3a.Multipart.Size=128M: es un aumento en comparación con los 64 millones.
- Out-of-the-box optimizaciones que mejoran el rendimiento para acelerar los trabajos con el sistema MapReduce de archivos S3A.

Amazon EMR 7.3.0: notas de la versión de Hadoop

Amazon EMR 7.3.0: cambios en Hadoop

Tipo	Descripción
Nueva característica	HADOOP-19197: S3A Support Encryption Context AWS KMS
Nueva característica	HADOOP-18980: Reasignación del proveedor de credenciales S3A: hacer extensible
Mejora	HADOOP-18808: para añadir un método debug () LogExactlyOnce

Amazon EMR 7.3.0: características de Hadoop

Consulte la siguiente lista para ver las nuevas funciones de Hadoop en Amazon EMR 7.3.0.

- De forma predeterminada, los clústeres con el cifrado en tránsito activado mediante la configuración de seguridad ejecutarán TLS 1.3 para Hadoop YARN (Resource Manager, Node Manager, Timeline Server y WebApplicationProxy), HDFS (NameNode DataNode, Journal Node and DFSZKFailover Controller) y Map Reduce (puertos Server y Shuffle). JobHistory MapReduce

Amazon EMR 7.2.0: notas de la versión de Hadoop

Amazon EMR 7.2.0: cambios en Hadoop

Tipo	Descripción
Nueva característica	HADOOP-18850: S3A: habilita el cifrado de doble capa del lado del servidor con claves AWS KMS
Nueva característica	La compatibilidad de S3 Glacier lee objetos restaurados
Mejora	Escalar dinámicamente el máximo de conexiones del S3A en función de los núcleos al ejecutar los trabajos de Spark
Mejora	HADOOP-18797: Compatibilidad para escritura s simultáneas con el confirmador S3A Magic
Mejora	HADOOP-19047: InMemory Support Tracking of S3A Magic Commits
Mejora	Añada los métodos «Exportar», «Publicar» y «» en MagicV2 Committer AbortPublish
Mejora	Implemente la API en ReadFullyIntoBuffers FileSystem S3A

Tipo	Descripción
Corrección de errores	HADOOP-18793: El S3A StagingCommitter no limpia el directorio staging-uploads
Upgrade	HADOOP-19008: S3A: actualice la versión aws-sdk a la 2.21.41
Upgrade	HADOOP-18613 ZooKeeper : Actualización a la versión 3.8.3

Amazon EMR 7.2.0: características de Hadoop

Consulte la siguiente lista para ver las nuevas funciones de Hadoop en Amazon EMR 7.2.0.

- Compatibilidad con S3 Glacier para leer objetos restaurados. Con Amazon EMR versión 7.2.0 y versiones posteriores, puede leer los objetos Glacier restaurados desde la ubicación S3 de la tabla con el protocolo `S3A`. Esta operación de lectura ignora los archivos de Glacier archivados si aún se están restaurando.
- Admite escrituras simultáneas con el confirmador S3A Magic: Amazon EMR 7.2.0 incorpora la compatibilidad para escrituras simultáneas con el confirmador S3A Magic. Esta función se agregó a la comunidad Hadoop de OSS como parte de HADOOP-18797.

Amazon EMR 6.6.0: notas de la versión de Hadoop

Amazon EMR 6.6.0: cambios en Hadoop

Tipo	Descripción
Error	Se corrigieron los registros duplicados al leer archivos de texto. BZip2
Portabilidad con versiones anteriores	HADOOP-18136: comprobación del manejo de `FileUtils.unTar()` de los archivos .tar que faltan

Tipo	Descripción
Portabilidad con versiones anteriores	HADOOP-17627: adaptación a branch-3.2 HADOOP-17371, HADOOP-17621, HADOOP-17625 para actualizar Jetty a la versión 9.4.39
Portabilidad con versiones anteriores	HADOOP-17655: actualización de Jetty a la versión 9.4.40
Portabilidad con versiones anteriores	HADOOP-17796: actualización de Jetty a la versión 9.4.43
Portabilidad con versiones anteriores	HADOOP-17661: mvn versions:set no puede analizar pom.xml
Portabilidad con versiones anteriores	HADOOP-17236: aumento de la versión de snakeyaml a 1.26 para mitigar CVE-2017-18640
Portabilidad con versiones anteriores	HADOOP-16717: Elimina la dependencia de IsLog4JLogger en Log4 GenericsUtil jLoggerAdapter
Portabilidad con versiones anteriores	HADOOP-17633: Salta json-smart a 2.4.2 y 9.8 debido a nimbus-jose-jwt CVEs
Portabilidad con versiones anteriores	HADOOP-17844: actualización de JSON-smart a la versión 2.4.7
Portabilidad con versiones anteriores	HADOOP-17972: adaptación de HADOOP-17683 (actualización de commons-io a la versión 2.8.0) para branch-3.2
Portabilidad con versiones anteriores	HADOOP-16555: actualización de commons-compress a la versión 1.19
Portabilidad con versiones anteriores	HADOOP-17370: actualización de commons-compress a la versión 1.21

Tipo	Descripción
Portabilidad con versiones anteriores	HADOOP-17096: corrige el ZStandard desfase del búfer de entrada del compresor
Portabilidad con versiones anteriores	HADOOP-17112: no se permiten espacios en blanco en las rutas cuando se guardan archivos en s3a mediante un confirmador
Portabilidad con versiones anteriores	HADOOP-13500: sincronización de iteración del objeto de propiedades de configuración
Portabilidad con versiones anteriores	HDFS-14099: Descriptor de fotogramas desconocido al descomprimir varios fotogramas en Decompressor ZStandard
Portabilidad con versiones anteriores	HDFS-16410: análisis XML inseguro en OfflineEditsXmlLoader
Portabilidad con versiones anteriores	HDFS-14498: LeaseManager puede repetirse indefinidamente en el archivo cuya creación ha fallado
Portabilidad con versiones anteriores	HDFS-15290: Entrada NPE durante el inicio HttpServer NameNode
Portabilidad con versiones anteriores	HDFS-15293: relajación de la condición para aceptar un fsimage al recibir un punto de control
Portabilidad con versiones anteriores	HDFS-12979: debe cargarse después del punto de control StandbyNode FsImage ObserverNode
Portabilidad con versiones anteriores	YARN-10538: agregación de los nodos que se vuelven a poner en servicio a la lista de nodos actualizados devueltos al AM

Tipo	Descripción
Portabilidad con versiones anteriores	YARN-10472: Conecta el YARN-10314 (se utiliza solo con los tarros de clientes sombreados) a la ramificación-3.2 YarnClient NoClassDefFoundError WebSocketException
Portabilidad con versiones anteriores	YARN-9968: Public Localizer se está cerrando en `NodeManager` debido a `NullPointerException`
Portabilidad con versiones anteriores	YARN-10651: `CapacityScheduler` dejó de funcionar con `NPE` en `AbstractYarnScheduler.updateNodeResource()`
Portabilidad con versiones anteriores	YARN-9339: la métrica de aplicaciones pendientes es incorrecta después de mover la aplicación a una nueva cola
Portabilidad con versiones anteriores	YARN-10438: gestión de `containerId` nulo en `ClientRMService#getContainerReport()`
Portabilidad con versiones anteriores	YARN-7266: ATS 1.5 no se inicia si los archivos `RollingLevelDb` están dañados o faltan
Portabilidad con versiones anteriores	YARN-9063: ATS 1.5 no se inicia si los archivos `RollingLevelDb` están dañados o faltan
Portabilidad con versiones anteriores	YARN-9848: revierta YARN-4946 (RM no debería considerar que una solicitud está `COMPLETED` cuando la agrupación de registros no está en un estado terminal).

Apache HBase

HBasees una base de datos distribuida, no relacional y de código abierto desarrollada como parte del proyecto Hadoop de la Apache Software Foundation. HBase se ejecuta sobre el sistema de archivos distribuido Hadoop (HDFS) para proporcionar capacidades de bases de datos no relacionales para el ecosistema de Hadoop. HBase se incluye en la versión 4.6.0 y posteriores de Amazon EMR.

HBase funciona a la perfección con Hadoop, ya que comparte su sistema de archivos y sirve como entrada y salida directas al MapReduce marco y al motor de ejecución. HBase también se integra con Apache Hive, lo que permite realizar consultas similares a las de SQL sobre HBase tablas, combinarlas con tablas basadas en Hive y es compatible con la conectividad de bases de datos Java (JDBC). Para obtener más información HBase, consulte Apache y la documentación en el sitio web de Apache HBase. HBase Para ver un ejemplo de cómo usarlo HBase con Hive, consulte la entrada del blog sobre AWS big data Combine NoSQL y análisis masivamente paralelo con HBase Apache y Apache Hive en Amazon EMR.

Con HBase Amazon EMR, también puede hacer copias de seguridad de sus HBase datos directamente en Amazon Simple Storage Service (Amazon S3) y restaurarlos a partir de una copia de seguridad creada anteriormente al lanzar un clúster. HBase Amazon EMR ofrece opciones adicionales para integrarse con Amazon S3 para la persistencia de datos y la recuperación de desastres.

- HBase en Amazon S3: con Amazon EMR versión 5.2.0 y versiones posteriores, puede utilizarlo en Amazon HBase S3 para almacenar el directorio HBase raíz y los metadatos de un clúster directamente en Amazon S3. Posteriormente, puede iniciar un nuevo clúster, que apunte a la ubicación del directorio raíz en Amazon S3. Solo un clúster a la vez puede usar la HBase ubicación en Amazon S3, con la excepción de un clúster de lectura y réplica. Para obtener más información, consulte HBase en Amazon S3 (modo de almacenamiento de Amazon S3).
- HBase réplicas de lectura: la versión 5.7.0 y posteriores de Amazon EMR, incluida en Amazon HBase S3, admite clústeres de réplica y lectura. Un clúster de réplicas de lectura proporciona acceso de solo lectura a los archivos de almacén y metadatos de un clúster primario para operaciones de solo lectura. Para obtener más información, consulte Uso de un clúster de réplicas de lectura.
- HBase Instantáneas: como alternativa a HBase Amazon S3, con la versión 4.0 de EMR y versiones posteriores, puede crear instantáneas de HBase sus datos directamente en Amazon S3

y, a continuación, recuperarlos con las instantáneas. Para obtener más información, consulte Uso de instantáneas HBase .

> **Important**
>
> Para el escalado de HBase clústeres de Amazon EMR, no recomendamos utilizar el escalado gestionado ni el escalado con políticas personalizadas con clústeres. HBase

En la siguiente tabla se muestra la versión HBase incluida en la última versión de la serie Amazon EMR 7.x, junto con los componentes con los que se instala Amazon EMR. HBase

Para ver la versión de los componentes que se incluyen HBase en esta versión, consulte Versiones de componentes de la versión 7.9.0.

HBase información sobre la versión de emr-7.9.0

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.9.0	HBase 2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

En la siguiente tabla se muestra la versión HBase incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes con los que se instala Amazon EMR. HBase

Para ver la versión de los componentes que se incluyen HBase en esta versión, consulte Versiones de componentes de la versión 6.15.0.

HBase información sobre la versión de emr-6.15.0

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.15.0	HBase 2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

> **Note**
>
> Apache HBase HBCK2 es una herramienta operativa independiente para reparar HBase regiones y tablas del sistema. En la versión 6.1.0 y posteriores de Amazon EMR, el archivo hbase-hbck2.jar se proporciona en `/usr/lib/hbase-operator-tools/` en el nodo principal. Para obtener más información sobre cómo crear y utilizar la herramienta, consulte HBaseHBCK2.

En la siguiente tabla se muestra la versión HBase incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes con los que se instala Amazon EMR. HBase

Para ver la versión de los componentes que se incluyen HBase en esta versión, consulte Versiones de componentes de la versión 5.36.2.

HBase información sobre la versión de emr-5.36.2

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.36.2	HBase 1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Temas

- Crear un clúster con HBase
- HBase en Amazon S3 (modo de almacenamiento de Amazon S3)
- Write Ahead Logs (WAL) para Amazon EMR
- Uso del HBase shell
- Acceda a HBase las tablas con Hive
- Uso de instantáneas HBase
- Configuración HBase
- Vea la interfaz de HBase usuario

- Ver archivos de HBase registro
- Supervise HBase con Ganglia
- Supervisión de EMR con Amazon HBase CloudWatch
- ¿Estás migrando desde versiones anteriores HBase
- HBase historial de versiones

Crear un clúster con HBase

Los procedimientos de esta sección abarcan los aspectos básicos del lanzamiento de un clúster mediante el AWS Management Console y el AWS CLI. Para obtener información detallada sobre cómo planificar, configurar y lanzar clústeres de Amazon EMR, consulte Planificar y configurar clústeres en la Guía de administración de Amazon EMR.

Crear un clúster HBase mediante la consola

Para ver los pasos rápidos para lanzar clústeres con la consola, consulte Introducción a Amazon EMR en la Guía de administración de Amazon EMR.

Para lanzar un clúster con HBase instalado, utilice la consola

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Create cluster (Crear clúster) y Go to advanced options (Ir a las opciones avanzadas).
3. En Software Configuration (Configuración de software), elija la Amazon Release Version (Versión de Amazon) 4.6.0 o posterior (recomendamos la versión más reciente). Elija HBasey otras aplicaciones según desee.
4. Con Amazon EMR versión 5.2.0 y versiones posteriores, en Configuración de HBase almacenamiento, seleccione HDFS o S3. Para obtener más información, consulte HBase en Amazon S3 (modo de almacenamiento de Amazon S3).
5. Seleccione las demás opciones que necesite y, a continuación, elija Create cluster (Crear clúster).

Creación de un clúster mediante el uso de HBase AWS CLI

Utilice el siguiente comando para crear un clúster con el comando HBase instalado:

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Si la usa HBase en Amazon S3, especifique la `--configurations` opción con una referencia a un objeto de configuración JSON. El objeto de configuración debe contener una `hbase-site` clasificación que especifique la ubicación en Amazon S3 donde se almacenan HBase los datos mediante la `hbase.rootdir` propiedad. También debe contener una clasificación `hbase`, que especifica `s3` utilizando la propiedad `hbase.emr.storageMode`. El siguiente ejemplo muestra un fragmento de JSON con estas opciones de configuración.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3"
        }
    }
]
```

Para obtener más información HBase sobre Amazon S3, consulteHBase en Amazon S3 (modo de almacenamiento de Amazon S3). Para obtener más información sobre las clasificaciones, consulte Configuración de aplicaciones.

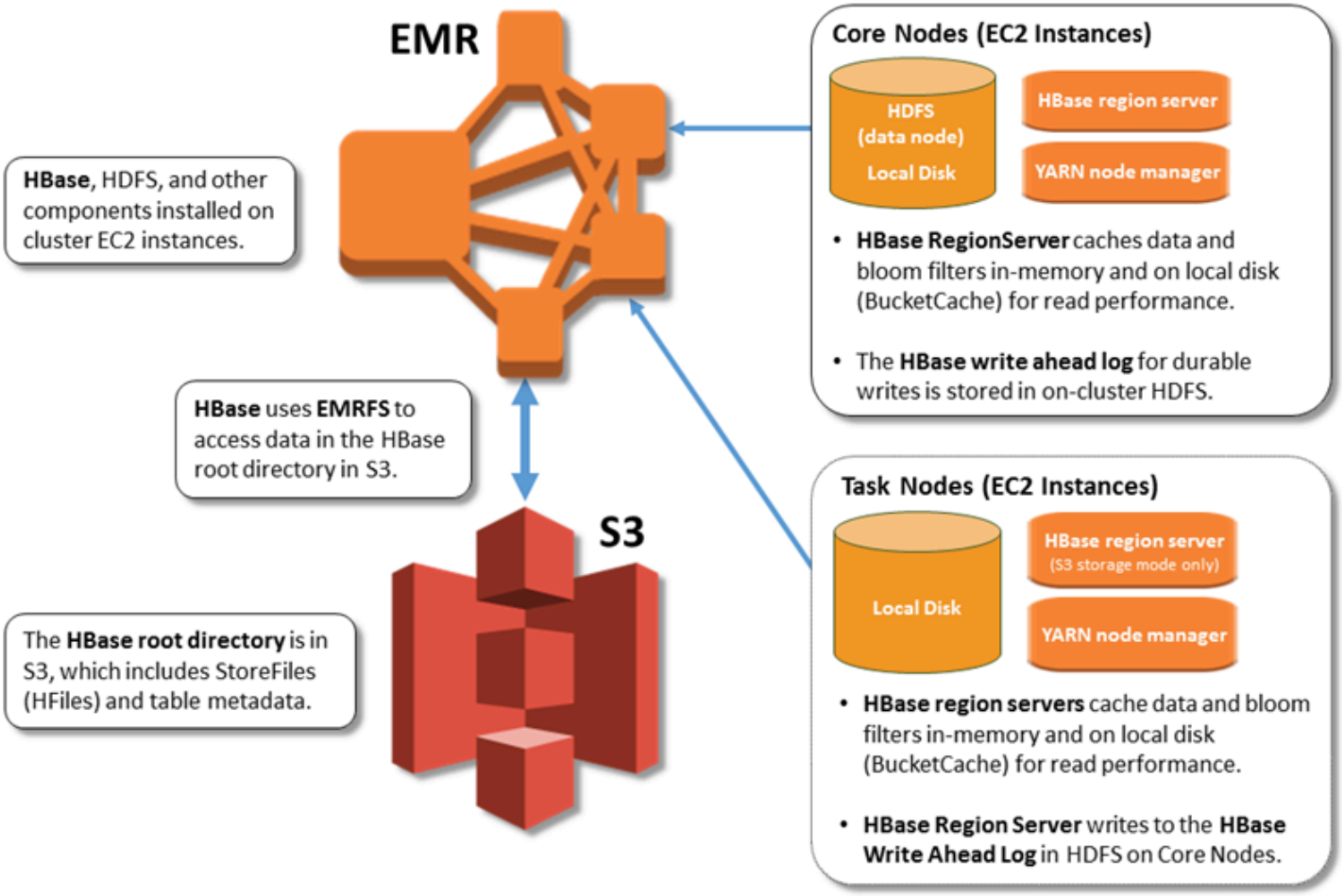
HBase en Amazon S3 (modo de almacenamiento de Amazon S3)

Si ejecuta Amazon EMR HBase en la versión 5.2.0 o posterior, puede habilitarlo en Amazon HBase S3, lo que ofrece las siguientes ventajas:

- El directorio HBase raíz se almacena en Amazon S3, incluidos los archivos de la HBase tienda y los metadatos de las tablas. Estos datos son persistentes fuera del clúster, están disponibles en todas las zonas de EC2 disponibilidad de Amazon y no es necesario recuperarlos mediante instantáneas u otros métodos.
- Con los archivos de almacén en Amazon S3, puede dimensionar el clúster de Amazon EMR en función de sus requisitos informáticos en lugar de sus requisitos de datos, con una replicación 3x en HDFS.
- Mediante la versión 5.7.0 de Amazon EMR o posteriores, puede configurar un clúster de réplicas de lectura, lo que le permite mantener copias de solo lectura de los datos en Amazon S3. Puede acceder a los datos desde el clúster de réplicas de lectura para realizar operaciones de lectura de forma simultánea y en caso de que el clúster principal deje de estar disponible.
- En las versiones 6.2.0 a 7.3.0 de Amazon EMR, el HFile seguimiento persistente utiliza una tabla HBase del sistema llamada `hbase:storefile` para rastrear directamente las HFile rutas utilizadas para las operaciones de lectura. Esta característica está habilitada de forma predeterminada y no requiere ninguna migración manual. En las versiones posteriores a la 7.3.0, las HFile rutas se rastrean mediante un rastreador de archivos, que almacena las HFile rutas directamente en un metaarchivo, dentro del directorio de la tienda.

> **Note**
>
> Los usuarios que utilizan una versión de Amazon EMR anterior a la 7.4.0 y van a migrar a EMR-7.4.0 y versiones posteriores, consulte Migración desde HBase versiones anteriores y siga la documentación de actualización disponible para garantizar una transición fluida.

La siguiente ilustración muestra los HBase componentes relevantes para HBase Amazon S3.



Habilitación HBase en Amazon S3

Puede activarlo HBase en Amazon S3 mediante la consola de Amazon EMR AWS CLI, la o la API de Amazon EMR. La configuración es una opción durante la creación del clúster. Al utilizar la consola, elija la configuración mediante las Advanced options (Opciones avanzadas). Al utilizar la AWS CLI, utilice la opción `--configurations` para proporcionar un objeto de configuración JSON. Las propiedades del objeto de configuración especifican el modo de almacenamiento y la ubicación del directorio raíz en Amazon S3. La ubicación de Amazon S3 que especifique debe estar en la misma región que el clúster de Amazon EMR. Solo un clúster activo a la vez puede utilizar el mismo directorio HBase raíz en Amazon S3. Para ver los pasos de la consola y un ejemplo detallado de creación de clústeres utilizando el AWS CLI, consulte. Crear un clúster con HBase Un objeto de configuración de ejemplo se muestra en el siguiente fragmento de JSON.

```
{
  "Classification": "hbase-site",
```

```
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir"}
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3"
  }
}
```

> **Note**
>
> Si utiliza un bucket de Amazon S3 como `rootdir` formulario HBase, debe añadir una barra al final del URI de Amazon S3. Por ejemplo, debe utilizar `"hbase.rootdir: s3://amzn-s3-demo-bucket/"`, en lugar de `"hbase.rootdir: s3://amzn-s3-demo-bucket"`, para evitar problemas.

Uso de un clúster de réplicas de lectura

Después de configurar un clúster principal con HBase Amazon S3, puede crear y configurar un clúster de réplica y lectura que proporcione acceso de solo lectura a los mismos datos que el clúster principal. Esto resulta útil cuando se necesita acceso simultáneo a datos de consulta o acceso ininterrumpido si el clúster principal deja de estar disponible. La característica de réplica de lectura está disponible en la versión 5.7.0 de Amazon EMR y posteriores.

El clúster principal y el clúster de réplicas de lectura se configuran de la misma forma con una importante diferencia. Ambos apuntan a la misma ubicación `hbase.rootdir`. Sin embargo, la clasificación `hbase` del clúster de réplicas de lectura incluye la propiedad `"hbase.emr.readreplica.enabled":"true"`.

El clúster de réplica y lectura está diseñado para operaciones de solo lectura y no se debe realizar ninguna acción manual de compactación o escritura en él. Para las versiones de Amazon EMR anteriores a la 7.4.0, se recomienda deshabilitar la compactación en el clúster de lectura-réplica al habilitar la función de lectura-réplica. Esta precaución es necesaria porque, con la función de HFile seguimiento persistente habilitada en el clúster principal, es posible que el clúster de lectura-réplica compacte las tablas del sistema, lo que podría provocar una en el clúster principal.

FileNotFoundException Al deshabilitar la compactación en el clúster de lectura-réplica, se evitan las incoherencias de datos entre el clúster principal y el clúster de réplica de lectura.

Por ejemplo, dada la clasificación JSON del clúster principal, tal como se muestra anteriormente en el tema, la configuración de un clúster de lectura-réplica para las versiones de EMR anteriores a la 7.4.0 es la siguiente:

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir",
    "hbase.regionserver.compaction.enabled": "false"
  }
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3",
  "hbase.emr.readreplica.enabled":"true"
  }
}
```

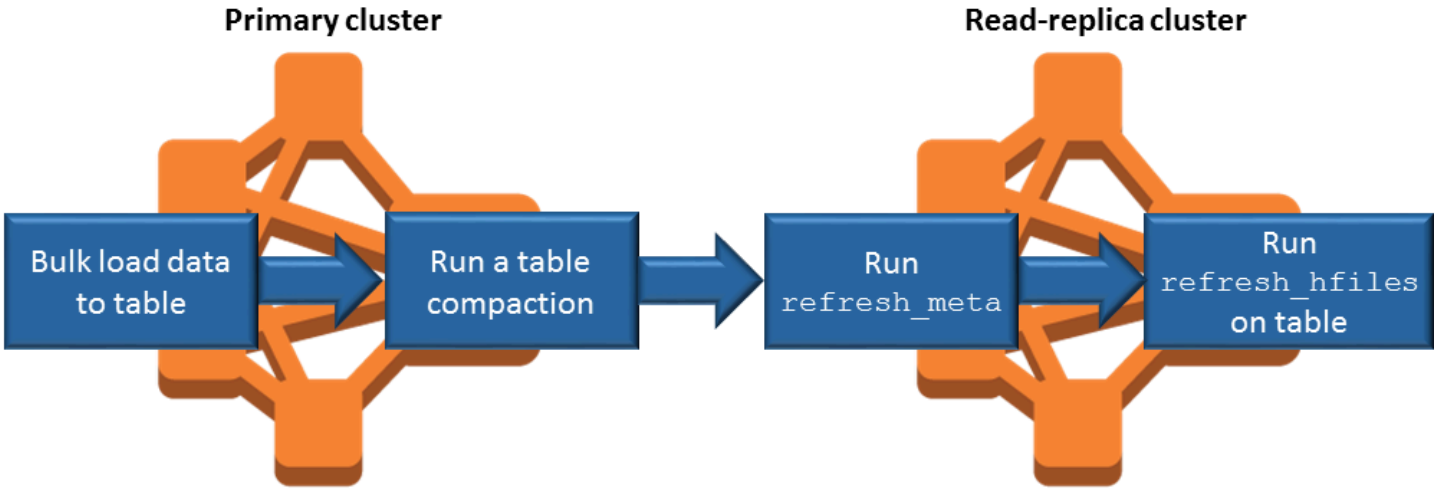
Para las versiones de Amazon EMR posteriores a la 7.3.0, ahora utilizamos la Seguimiento de archivos de la tienda función, por lo que no es necesario deshabilitar las compactaciones.

Sincronización de la réplica de lectura al agregar datos

Como la réplica de lectura utiliza HBase StoreFiles los metadatos que el clúster principal escribe en Amazon S3, la réplica de lectura solo es tan actualizada como el almacén de datos de Amazon S3. Las siguientes directrices pueden ayudarle a minimizar el retardo entre el clúster principal y la réplica de lectura al escribir datos.

- Cargue los datos de forma masiva en el clúster principal siempre que sea posible. Para obtener más información, consulte la documentación sobre la carga masiva en Apache. HBase
- Debe llevarse a cabo un vaciado que escriba archivos de almacén en Amazon S3 tan pronto como sea posible después de agregar los datos. Realice el vaciado manualmente o bien ajuste la configuración de vaciado para minimizar el retardo.

- Si las compactaciones se pudieran ejecutar automáticamente, ejecute una compactación manual para evitar incoherencias cuando se activen las compactaciones.
- En el clúster de lectura-réplica, ejecute el comando cuando se modifique algún metadato (por ejemplo, cuando se dividan o compacten HBase regiones, o cuando se agreguen o eliminen tablas). `refresh_meta`
- En el clúster de réplicas de lectura, ejecute el comando `refresh_hfiles` cuando se añadan registros o se cambien en una tabla.



Seguimiento persistente HFile

HFile El seguimiento persistente utiliza una tabla HBase del sistema llamada `hbase:storefile` para rastrear directamente las HFile rutas utilizadas para las operaciones de lectura. Las nuevas HFile rutas se añaden a la tabla a medida que se añaden datos adicionales HBase. Esto elimina las operaciones de cambio de nombre como mecanismo de confirmación en HBase las rutas de escritura críticas y mejora el tiempo de recuperación al abrir una HBase región al leer la tabla del `hbase:storefile` sistema en lugar de la lista de directorios del sistema de archivos. Esta función está habilitada de forma predeterminada en las versiones 6.2.0 a 7.3.0 de Amazon EMR y no requiere ningún paso de migración manual.

> **Note**
>
> El HFile seguimiento persistente mediante la tabla del HBase sistema de archivos de almacenamiento no admite la HBase función de replicación regional. Para obtener más

información sobre HBase la replicación regional, consulte Lecturas de alta disponibilidad coherentes con el cronograma.

Desactivar el seguimiento persistente HFile

El HFile seguimiento persistente está activado de forma predeterminada a partir de la versión 6.2.0 de Amazon EMR. Para deshabilitar el HFile seguimiento persistente, especifique la siguiente anulación de configuración al lanzar un clúster:

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.storefile.tracking.persist.enabled":"false",

 "hbase.hstore.engine.class":"org.apache.hadoop.hbase.regionserver.DefaultStoreEngine"
  }
}
```

Note

Al volver a configurar el clúster de Amazon EMR, se deben actualizar todos los grupos de instancias.

Sincronización manual de la tabla de archivos de almacenamiento

La tabla de archivos de almacenamiento se mantiene actualizada a medida que se crean nuevas HFile instancias. Sin embargo, si la tabla de archivos de almacenamiento no está sincronizada con los archivos de datos por algún motivo, se pueden usar los siguientes comandos para sincronizar los datos manualmente:

Sincronice la tabla de archivos de almacenamiento en una región en línea:

```
hbase org.apache.hadoop.hbase.client.example.RefreshHFilesClient <table>
```

Sincronice la tabla de archivos de almacenamiento en una región sin conexión:

- Elimine la tabla de archivos de almacenamiento znode.

```
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
```

```
[<tableName>, hbase:namespace]
# The TableName exists in the list
echo "delete /hbase/storefile/loaded/<tableName>" | sudo -u hbase hbase zkcli
# Delete the Table ZNode
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
[hbase:namespace]
```

- Asigne la región (ejecute en el intérprete de comandos de HBase).

```
hbase cli> assign '<region name>'
```

- Si se produce un error en la asignación.

```
hbase cli> disable '<table name>'
hbase cli> enable '<table name>'
```

Escalar la tabla de archivos de almacenamiento

La tabla de archivos de almacenamiento se divide en cuatro regiones de forma predeterminada. Si la tabla de archivos de almacenamiento aún tiene una gran carga de escritura, se puede dividir aún más manualmente.

Para dividir una región activa específica, utilice el siguiente comando (ejecute en el intérprete de comandos de HBase).

```
hbase cli> split '<region name>'
```

Para dividir la tabla, utilice el siguiente comando (ejecute en el intérprete de comandos de HBase).

```
hbase cli> split 'hbase:storefile'
```

Seguimiento de archivos de la tienda

De forma predeterminada, utilizamos la FileBasedStoreFileTrackerimplementación. Esta implementación crea nuevos archivos directamente en el directorio de la tienda, lo que evita la necesidad de realizar operaciones de cambio de nombre. Mantiene una lista de las instancias de hfile confirmadas en la memoria, respaldada por los metaarchivos de cada directorio de la tienda. Cada vez que se confirma un nuevo archivo hfile, se actualiza la lista de archivos rastreados en el almacén en cuestión y se escribe un nuevo metaarchivo con el contenido de la lista y descartando

el metarchivo anterior, que contiene una lista obsoleta. Puede encontrar más información sobre el seguimiento de archivos de almacén en la guía de HBase referencia de Apache en el apartado Seguimiento de archivos de almacén.

La implementación del FileBasedStoreFile rastreador está habilitada de forma predeterminada, a partir de la versión 7.4.0 de Amazon EMR:

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.FileBasedStoreFileTracker"
}
```

Para deshabilitar la FileBasedStoreFileTracker implementación, especifique la siguiente anulación de configuración al lanzar un clúster:

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

> **Note**
>
> Al volver a configurar el clúster de Amazon EMR, se deben actualizar todos los grupos de instancias.

Consideraciones operativas

HBase los servidores de región se utilizan BlockCache para almacenar las lecturas de datos en la memoria y BucketCache para almacenar las lecturas de datos en el disco local. Además, los servidores de región suelen almacenar MemStore las escrituras de datos en la memoria y utilizan los registros de escritura anticipada para almacenar las escrituras de datos en HDFS antes de que los datos se escriban en Amazon HBase StoreFiles S3. El rendimiento de lectura del clúster se refiere a la frecuencia con la que un registro puede recuperarse en memoria o en las cachés de disco. Si se pierde la memoria caché, el registro se lee desde Amazon S3, que tiene una latencia y una

desviación estándar significativamente más altas que la lectura desde HDFS. StoreFile Además, las velocidades de solicitud máximas en Amazon S3 son más bajas que las que se pueden conseguir en la caché local, por lo que el almacenamiento de datos en caché podría ser importante para cargas de trabajo con lectura intensiva. Para obtener más información, consulte Optimizar el rendimiento de Amazon S3 en la Guía del usuario de Amazon Simple Storage Service.

Para mejorar el rendimiento, le recomendamos que almacene en caché la mayor cantidad posible de su conjunto de datos en el almacenamiento de EC2 instancias. Como BucketCache utiliza el almacenamiento de EC2 instancias del servidor de la región, puede elegir un tipo de EC2 instancia con un almacén de instancias suficiente y añadir el almacenamiento de Amazon EBS para acomodar el tamaño de caché requerido. También puede aumentar el BucketCache tamaño de los almacenes de instancias adjuntos y los volúmenes de EBS mediante la `hbase.bucketcache.size` propiedad. La configuración predeterminada es 8 192 MB.

En el caso de las escrituras, la MemStore frecuencia de las descargas y el número de descargas StoreFiles presentes durante las compactaciones menores y mayores pueden contribuir considerablemente a aumentar los tiempos de respuesta de los servidores regionales. Para obtener un rendimiento óptimo, considere la posibilidad de aumentar el tamaño del multiplicador de HRegion bloques y de MemStore descarga, lo que aumentará el tiempo transcurrido entre las compactaciones principales, pero también aumentará el retraso en la coherencia si utiliza una réplica de lectura. En algunos casos, es posible obtener un mejor rendimiento mediante tamaños de bloque de archivo más grandes (pero inferiores a 5 GB) para activar la funcionalidad de carga multiparte de Amazon S3 en EMRFS. El tamaño de bloque predeterminado de Amazon EMR es de 128 MB. Para obtener más información, consulte Configuración de HDFS. No son frecuentes los clientes que superen un tamaño de bloque de 1 GB al realizar un análisis comparativo del rendimiento con vaciados y compactaciones. Además, HBase las compactaciones y los servidores regionales funcionan de forma óptima cuando es necesario compactar menos. StoreFiles

Las tablas pueden tardar mucho en eliminarse en Amazon S3, ya que es necesario cambiar de nombre directorios grandes. Considere la posibilidad de deshabilitar las tablas en lugar de eliminarlas.

Existe un proceso HBase más limpio que limpia los archivos WAL antiguos y almacena los archivos. En la versión 5.17.0 de Amazon EMR y posteriores, el limpiador está habilitado de forma global y se pueden utilizar las siguientes propiedades de configuración para controlar su comportamiento.

Propiedad de configuración	Valor predeterminado	Descripción

Propiedad de configuración	Valor predeterminado	Descripción
`hbase.regionserver.hfilecleaner.large.thread.count`	1	El número de subprocesos asignados a la limpieza ha caducado mucho. HFiles
`hbase.regionserver.hfilecleaner.small.thread.count`	1	El número de subproces os asignados a la limpieza expirado es pequeño HFiles.
`hbase.cleaner.scan.dir.concurrent.size`	Se establece en la cuarta parte de los núcleos disponibl es.	El número de subprocesos para escanear los WALs directorios antiguos.
`hbase.oldwals.cleaner.thread.size`	2	El número de subprocesos que se van a limpiar WALs en el WALs directorio antiguo.

En Amazon EMR 5.17.0 y versiones anteriores, el funcionamiento del limpiador puede afectar al rendimiento de las consultas cuando se ejecutan cargas de trabajo pesadas, por lo que le recomendamos que únicamente habilite el limpiador fuera de las horas punta. El limpiador tiene los siguientes comandos de HBase shell:

- `cleaner_chore_enabled` consulta si el limpiador está habilitado.
- `cleaner_chore_run` ejecuta manualmente el limpiador para eliminar archivos.
- `cleaner_chore_switch` habilita o deshabilita el limpiador y devuelve el estado anterior del limpiador. Por ejemplo, `cleaner_chore_switch true` habilita el limpiador.

Propiedades para ajustar HBase el rendimiento en Amazon S3

Los siguientes parámetros se pueden ajustar para ajustar el rendimiento de la carga de trabajo cuando se utiliza HBase en Amazon S3.

Propiedad de configuración	Valor predeterminado	Descripción
`hbase.bucketcache.size`	8 192	La cantidad de espacio en disco, en MB, reservado en el servidor regional, los almacenes de EC2 instancias de Amazon y los volúmenes de EBS para el BucketCache almacenamiento. La configuración se aplica a todas las instancias de servidor de región. BucketCache Los tamaños más grandes suelen corresponder a un rendimiento mejorado
`hbase.hregion.memstore.flush.size`	134217728	El límite de datos, en bytes, a partir del cual el vaciado de memstore en Amazon S3 se activa.
`hbase.hregion.memstore.block.multiplier`	4	Multiplicador que determina el límite MemStore superior en el que se bloquean las actualizaciones. Si `hbase.hregion.memstore.flush.size` se MemStore supera el valor multiplicado por este valor, las actualizaciones se bloquean. MemStore Es posible que las descargas y la compactación desbloqueen las actualizaciones.
`hbase.hstore.blockingStoreFiles`	10	La cantidad máxima StoreFiles que puede haber en una

Propiedad de configuración	Valor predeterminado	Descripción
		tienda antes de que se bloqueen las actualizaciones.
`hbase.hregion.max.filesize`	10737418240	El tamaño máximo de una región antes de que la región se divida.

Cierre y restauración de un clúster sin pérdida de datos

Para cerrar un clúster de Amazon EMR sin perder datos que no se hayan escrito en Amazon S3, debe vaciar la MemStore memoria caché en Amazon S3 para escribir nuevos archivos de almacenamiento. En primer lugar, tendrá que deshabilitar todas las tablas. La siguiente configuración de paso puede utilizarse al añadir un paso al clúster. Para obtener más información, consulte Uso de pasos con la AWS CLI y la consola en la Guía de administración de Amazon EMR.

```
Name="Disable all tables",Jar="command-runner.jar",Args=["/bin/bash","/usr/lib/hbase/
bin/disable_all_tables.sh"]
```

De forma alternativa, puede ejecutar el siguiente comando bash directamente.

```
bash /usr/lib/hbase/bin/disable_all_tables.sh
```

Tras deshabilitar todas las tablas, vacíe la `hbase:meta` tabla con el HBase shell y el siguiente comando.

```
flush 'hbase:meta'
```

A continuación, puede ejecutar un script de shell incluido en el clúster de Amazon EMR para vaciar la caché. MemStore Puede agregarlo como un paso o ejecutarlo directamente a través de la AWS CLI en el clúster. El script deshabilita todas HBase las tablas, lo que provoca que MemStore el servidor de cada región pase a Amazon S3. Si el script se completa de forma satisfactoria, los datos persisten en Amazon S3 y el clúster puede terminarse.

Para reiniciar un clúster con los mismos HBase datos, especifique la misma ubicación de Amazon S3 que el clúster anterior, ya sea en la propiedad de `hbase.rootdir` configuración AWS Management Console o mediante ella.

Write Ahead Logs (WAL) para Amazon EMR

Con Amazon EMR 6.15 y versiones posteriores, puede escribir sus registros de HBase escritura anticipada (WAL) de Apache en el WAL de Amazon EMR. Con versiones anteriores de Amazon EMR, al crear un clúster con la opción HBase en Amazon S3, WAL es el único HBase componente de Apache que se almacena en el disco local para los clústeres, y puede almacenar otros componentes, como el directorio raíz, los archivos de almacenamiento (HFiles), los metadatos de las tablas y los datos en Amazon S3.

Puede usar Amazon EMR WAL para recuperar los datos que no se vaciaron en Amazon S3. Para realizar una copia de seguridad completa de sus HBase clústeres, opte por utilizar el servicio Amazon EMR WAL. Entre bastidores, `RegionServer` escribe sus registros de HBase escritura anticipada (WAL) en el WAL para Amazon EMR.

En caso de que su clúster o la zona de disponibilidad (AZ) estén en mal estado o no estén disponibles, puede crear un clúster nuevo, dirigirlo al mismo directorio raíz de S3 y al espacio de trabajo Amazon EMR WAL y recuperar automáticamente los datos en WAL en unos minutos. Para obtener más información, consulte Restauración desde Amazon EMR WAL.

A partir de las versiones 7.3.0 y posteriores de Amazon EMR, Amazon EMR crea varios EMR para cada servidor y agrupa varias regiones HBase en una sola WAL de Amazon WALs EMR. De este modo, se mejora el HBase WAL de Apache para mejorar la utilización de los registros y optimizar los costes. Para configurar el número de instancias WAL de Amazon EMR por instancia HBase `RegionServer`, utilice el parámetro. `hbase.wal.regiongrouping.numgroups` Este parámetro está establecido en 2 de forma predeterminada. Hay dos tablas del sistema que no se incluyen en ningún grupo WAL: meta y masterstore. Estas tablas siempre usan su propia tabla individual WALs.

Si ejecuta una versión anterior a Amazon EMR 7.3.0, le recomendamos que deshabilite manualmente las tablas del HBase clúster anterior para asegurarse de que todos los datos de la WAL de Amazon EMR se vacíen en Amazon S3. A continuación, elimine la Amazon EMR WAL anterior, finalice el clúster anterior y configure un clúster nuevo que ejecute la versión más reciente. Si tiene problemas y no puede deshabilitar las tablas del clúster anterior, puede finalizar directamente el clúster anterior y establecer `emr.wal.multiplex.migrate` en `true` en el nuevo clúster. Si se establece en true, HBase intentará reproducir los datos de las instancias WAL de Amazon EMR antiguas HBase durante la inicialización de la región y eliminará los antiguos después de la WALs reproducción. Este proceso de reproducción implica costes adicionales de lectura. Tras la migración, se recomienda configurar el clúster y establecer `emr.wal.multiplex.migrate` en `false`. Como alternativa, puede eliminar el parámetro para acelerar la inicialización de la región. HBase

> **Note**
>
> Amazon EMR WAL elimina los datos después HBase de vaciarlos. Si HBase no vacía los datos, Amazon EMR WAL los conserva durante un máximo de 30 días. Transcurridos 30 días, Amazon EMR WAL elimina automáticamente los datos. Amazon EMR conserva las instancias de WAL durante un máximo de 30 días a partir de la finalización de un clúster de EMR. Sin embargo, si lanza un nuevo clúster habilitado para WAL desde el mismo directorio raíz de S3 dentro de esos 30 días, Amazon EMR no eliminará ninguna de las instancias de WAL del clúster anterior. Para obtener más información, consulte Restauración desde Amazon EMR WAL.

En las siguientes secciones se describe cómo configurar y utilizar Amazon EMR WAL con su clúster de EMR HBase habilitado.

Temas

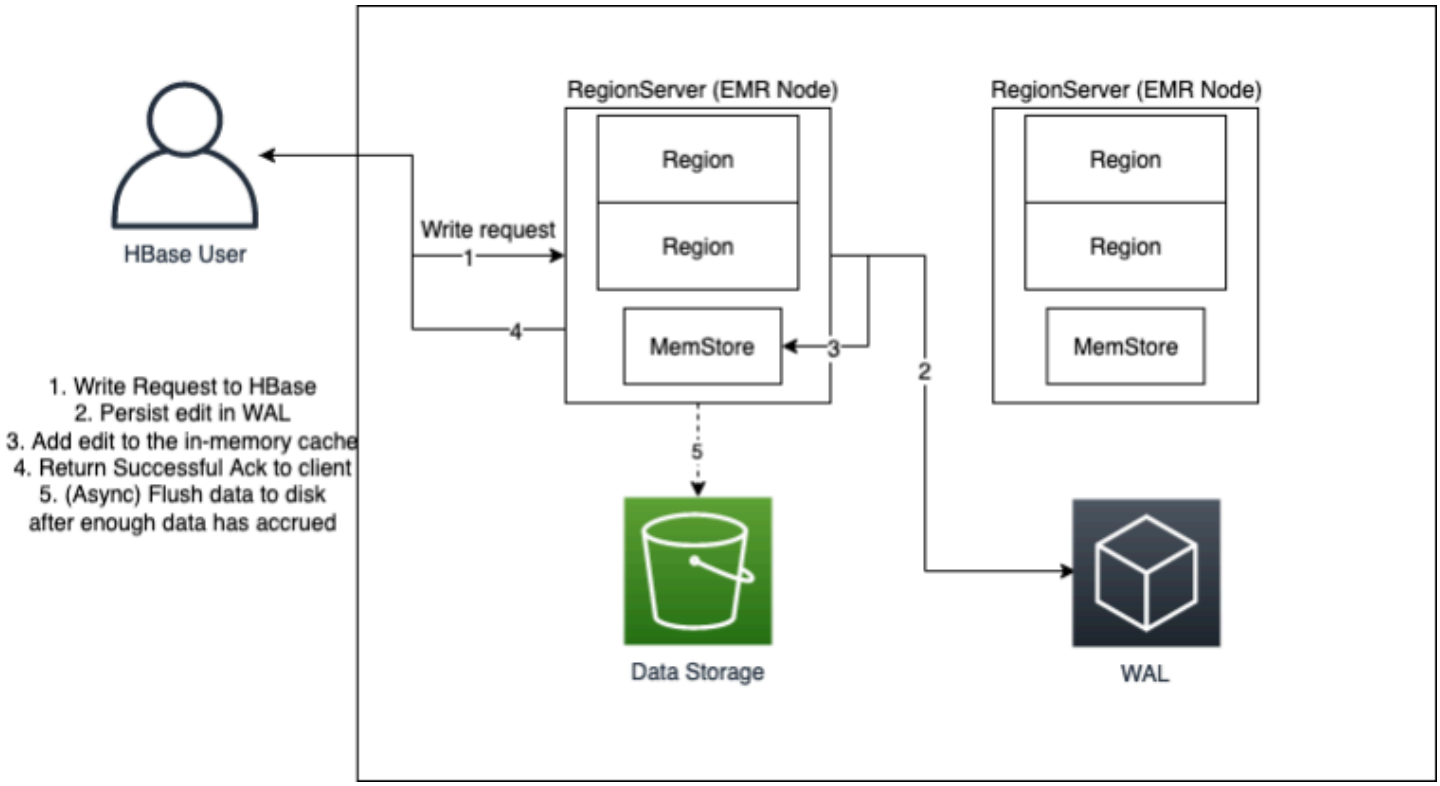
- Espacios de trabajo de Amazon EMR WAL
- Permisos necesarios para Amazon EMR WAL
- Habilitación de Amazon EMR WAL
- Restauración desde Amazon EMR WAL
- Uso de configuraciones de seguridad con Amazon EMR WAL
- Acceda a Amazon EMR WAL a través de AWS PrivateLink
- Descripción de los precios y las métricas de Amazon EMR WAL
- Etiquetado de espacios de trabajo de WAL
- Replicación entre clústeres EMR WAL
- Consideraciones y regiones para Amazon EMR WAL
- Referencia de la (EMRWAL) CLI de Amazon EMR WAL

Espacios de trabajo de Amazon EMR WAL

Amazon EMR WAL añade el concepto de espacios de trabajo de WAL. Un espacio de trabajo WAL es un contenedor lógico de. WALs Cada registro de escritura anticipada de Amazon EMR WAL está encapsulado en un espacio de trabajo de WAL. Un clúster de EMR escribe WALs exactamente en un espacio de trabajo de WAL que se configura al lanzar el clúster, o en el `defaultWALworkspace` si

no se especifica un espacio de trabajo. Los espacios de trabajo de WAL no están relacionados con ninguna HBase terminología existente, como los espacios de nombres.

Puede usar los espacios de trabajo de WAL para reducir el alcance de los permisos de la IAM de Amazon EMR WAL e incluir solo los espacios de trabajo a los que el clúster necesita acceder. También puede etiquetar su espacio de trabajo de WAL para el control de acceso basado en etiquetas. Para obtener más información acerca del etiquetado, consulte Etiquetado de espacios de trabajo de WAL:



Permisos necesarios para Amazon EMR WAL

Para que el clúster se conecte a Amazon EMR WAL, el perfil de instancia del clúster requiere determinados permisos de IAM:

- Amazon EMR utiliza el rol vinculado a servicios AWSServiceRoleForEMRWAL para recuperar el estado de un clúster. Amazon EMR crea automáticamente este rol vinculado al servicio cuando crea un espacio de trabajo de WAL, o HBase creará el rol vinculado al servicio cuando configure un espacio de trabajo para Amazon EMR WAL y el rol vinculado al servicio aún no existe.

  Antes de poder habilitar Amazon EMR WAL para un clúster, debe configurar los permisos para permitir la creación automática del rol vinculado a servicios AWSServiceRoleForEMRWAL. Para

obtener más información y un ejemplo de instrucción que añada esta capacidad, consulte Uso de roles vinculados a servicios para el registro de escritura anticipada.

- Dado que Amazon EMR WAL usa HBase Write Ahead Log (WAL), sus clústeres deben usar HBase WAL. Los siguientes son los permisos de IAM mínimos que necesita para ejecutarse. HBase Añada estos a la política de permisos para su perfil de instancia:

```
emrwal:DeleteWal
emrwal:CreateWal
emrwal:CreateWorkspace
emrwal:AppendEdit
emrwal:ReplayEdits
emrwal:GetCurrentWalTime
emrwal:CompleteWalFlush
emrwal:ListWALs
emrwal:DescribeWAL
emrwal:TrimWAL
emrwal:ArchiveWAL
emrwal:ArchiveWALCheckPoint
```

> **Note**
>
> Si limita el ámbito de los permisos de Amazon EMR WAL a solo al conjunto mínimo, algunos comandos de la EMRWALCLI no tendrán los permisos necesarios para ejecutarse.

Habilitación de Amazon EMR WAL

Siga los pasos siguientes para habilitar la escritura en Amazon EMR WAL al crear un clúster con AWS Command Line Interface.

> **Note**
>
> No puede habilitar Amazon EMR WAL para un clúster que ya se está ejecutando ni puede lanzar dos clústeres con el mismo directorio raíz de S3. Para obtener más información, consulte Consideraciones y regiones para Amazon EMR WAL.

1. Antes de poder crear un clúster habilitado para Amazon EMR WAL, debe añadir los permisos necesarios al perfil de instancia que planea usar con el clúster. Para obtener más información, consulte Permisos necesarios para Amazon EMR WAL.
2. Creación de un clúster desde la AWS CLI. Utilice la opción `--configurations` para proporcionar un objeto de configuración JSON que especifique la propiedad`hbase.emr.wal.enabled`, como se muestra en el siguiente ejemplo.

   - Las propiedades del objeto de configuración especifican el modo de almacenamiento y la ubicación del directorio raíz en Amazon S3. La ubicación de Amazon S3 que especifique debe estar en la misma región que su clúster de EMR, pero solo un clúster activo puede utilizar el mismo directorio HBase raíz en S3 a la vez.
   - Creación de un clúster con la configuración de grupos de instancia No puede usar Amazon EMR WAL con la configuración de las flotas de instancias. Para obtener más información sobre la creación de clústeres con grupos de instancias, consulte Configuración de grupos de instancias uniformes en la Guía de administración de Amazon EMR.
   - Para ver los pasos de la consola para crear un clúster y un `create-cluster` ejemplo detallado en el que se utiliza AWS CLI, consulte Crear un clúster con HBase.
3. Para habilitar WAL en el nuevo clúster, establezca la propiedad `hbase.emr.wal.enabled` en `true`. El siguiente comando contiene un fragmento de código JSON con un objeto de configuración de ejemplo.

```
aws emr create-cluster --name "hbasewal" --release-label emr-6.x.y \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m6i.xlarge --instance-count 1 --configurations hbase.json
$cat hbase.json
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
```

```
]
```

Cuando HBase esté en línea en el clúster recién creado, HBase escribirá automáticamente los datos de WAL en la WAL de Amazon EMR y utilizará la WAL de Amazon EMR con fines de recuperación.

Example 1: creación de un clúster de EMR que utilice Amazon EMR WAL

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Example 2: creación de un clúster de EMR con un espacio de trabajo de WAL personalizado

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Restauración desde Amazon EMR WAL

Como el Amazon EMR WAL de su clúster original se conserva durante 30 días, puede restaurar y reutilizar el WAL para un clúster recién creado dentro de ese período de 30 días. Cuando lanza un clúster nuevo desde el mismo directorio raíz de S3, Amazon EMR conserva las instancias del WAL del clúster anterior. Si finaliza este nuevo clúster, el reloj de 30 días se reiniciará a partir del momento en que lo finaliza.

Utilice el siguiente procedimiento para restaurar un WAL existente con un nuevo clúster. Este proceso supone que creó el clúster original con Amazon EMR WAL habilitado.

1. En un plazo de 30 días a partir de la creación de un clúster compatible con WAL, cree uno nuevo Región de AWS igual que el clúster original. El nuevo clúster puede estar en la misma AZ o en una AZ diferente dentro de la misma región en la que se creó el clúster original.

   Configuración de las propiedades del objeto para especificar el modo de almacenamiento y la ubicación del directorio raíz en Amazon S3. La ubicación de Amazon S3 que especifique debe estar en la misma región que su clúster de EMR, pero solo un clúster activo puede utilizar el mismo directorio HBase raíz en S3 a la vez.

   Para ver los pasos de la consola para crear un clúster y un `create-cluster` ejemplo detallado en el que se utiliza AWS CLI, consulte Crear un clúster con HBase.
2. Para usar el Amazon EMR WAL existente para el nuevo clúster, establezca la propiedad `hbase.emr.wal.enabled` en `true`. E siguiente fragmento de JSON muestra un objeto de configuración de ejemplo.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
```

```
    }
]
```

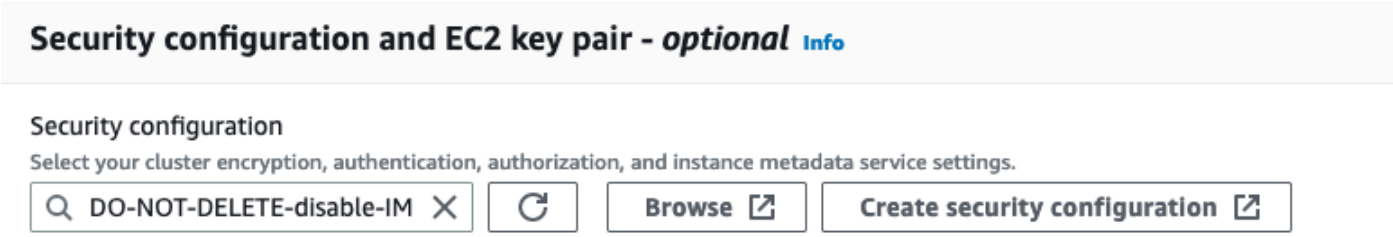
Uso de configuraciones de seguridad con Amazon EMR WAL

Amazon EMR cifra automáticamente los datos en tránsito entre el clúster y el servicio de Amazon EMR WAL, así como los datos en reposo en Amazon EMR WAL. Para obtener más información, consulte Cifrado en reposo de Amazon EMR WAL. También puede usar una configuración de seguridad para traer sus propias claves del servicio AWS Key Management Service (KMS) y cifrar los datos que almacena en Amazon EMR WAL.

Utilice uno de los siguientes métodos para seleccionar una configuración de seguridad al crear un clúster:

Console

Desde AWS Management Console, especifique la configuración en Configuración de seguridad y EC2 key pair.



CLI

Desde AWS CLI, defina el `--security-configuration` parámetro cuando utilice el comando create-cluster.

Para obtener más información, consulte Cifrado en reposo de Amazon EMR WAL y Uso de configuraciones de seguridad para definir la seguridad del clúster en la Guía de administración de Amazon EMR.

Para obtener más información relacionada con la seguridad sobre WAL, consulte Uso de roles vinculados a servicios para el registro anticipado.

Acceda a Amazon EMR WAL a través de AWS PrivateLink

Si desea mantener su conexión dentro de la AWS red, Amazon EMR WAL ofrece AWS PrivateLink soporte. Para configurarlo AWS PrivateLink, utilice AWS Management Console o AWS Command Line Interface (AWS CLI) para crear un punto final de VPC de interfaz que se conecte a Amazon EMR WAL. Para obtener más información, consulte Acceder a un AWS servicio mediante un punto final de VPC de interfaz en la AWS PrivateLink Guía.

Los pasos básicos son los siguientes:

1. Utilice la consola de Amazon VPC para crear un punto de conexión de VPC. Seleccione Puntos de conexión y, a continuación, Crear punto de conexión.
2. Mantenga la categoría de servicio como AWS servicios.
3. En la barra de búsqueda del panel de Servicios, escriba **`emrwal`** y, a continuación, seleccione el servicio etiquetado `com.amazonaws.`*`region`*`.emrwal.prod`.
4. Seleccione su VPC y guarde el punto de conexión. Asegúrese de conectar los mismos grupos de seguridad al punto de conexión de la VPC que conecta al clúster de EMR.
5. Si lo desea, ahora puede habilitar los nombres del host DNS privados para su nuevo punto de conexión. Establezca Habilitación de nombres de host DNS y Habilitación de soporte de DNS en `true` para su VPC. A continuación, seleccione su ID de punto de conexión, seleccione Editar la configuración de la VPC en el menú Acciones y habilite los nombres DNS privados.

   - Los nombres de host DNS privados del punto de conexión seguirán el formato `prod.emrwal.`*`region`*`.amazonaws.com`.
   - Si no habilita nombres de host DNS privados, Amazon VPC proporciona un nombre de punto de conexión de DNS que puede utilizar en el formato *`endpointID`*`.prod.emrwal.`*`region`*`.vpce.amazonaws.com`.
6. Para usar su AWS PrivateLink punto de conexión, modifique la `emr.wal.client.endpoint` configuración cuando cree su clúster habilitado para Amazon EMR WAL, como se muestra en el siguiente ejemplo:

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName",
            "emr.wal.client.endpoint": "https://prod.emrwal.region.amazonaws.com"
```

```
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

También puede utilizar la política VPCE para permitir o restringir el acceso a la WAL de Amazon APIs EMR. Para obtener más información, consulte Uso de políticas de punto de conexión para controlar el acceso a puntos de conexión de VPC en la Guía de AWS PrivateLink .

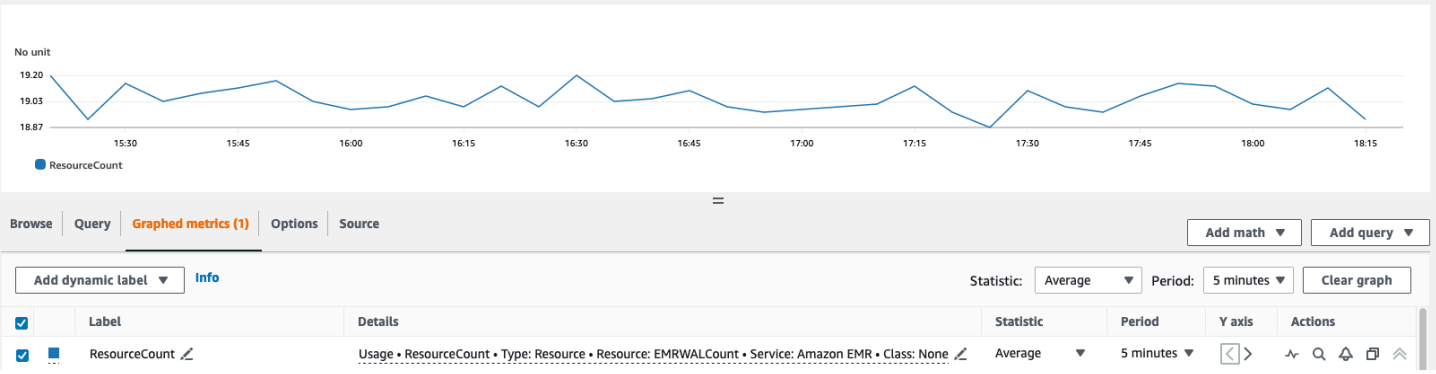
Descripción de los precios y las métricas de Amazon EMR WAL

Unidad de facturación de caracterí sticas principales	Detalles
EMR-WAL-Read-GiB	API calls to read data from your table are billed as ReadRequestGiB. This includes `Get` and `Scan` operations. Reads are charged based on the sizes of the read items. Amazon EMR bills at a minimum of 1 byte. For example, if you read a 1234.12 bytes item, you're charged for 1235 bytes. Reads are aggregated every hour for billing and shown as GiBs.
EMR-WAL-Write-GiB	API calls to write data from your table are billed as Write-GiB. This includes `Put` operations. Writes are charged based on the sizes of the written items. Amazon EMR bills at a minimum of 1 byte. For example, if you write a 1234.12 bytes item, you're charged for 1235 bytes. Writes are aggregated every hour for billing and shown as GiBs.
EMR-WAL-WALHours	El número WALs que almacenas en el servicio se factura como. `EMR-WAL-WALHours` Amazon EMR

Unidad de facturación de caracterí sticas principales	Detalles

	crea una WAL por región. HBase Por ejemplo, si crea 20 HBase tablas, incluidas las tablas del sistema, y cada tabla tiene dos HBase regiones, utilizará 28 800 horas de WAL, calculadas de la siguiente manera:
	`20 tables`
	`x 2 Regions per table`
	`x 1 WAL per Region`
	`x 30 days`
	`x 24 hours`
	`----------`
	`28,800 EMR-WAL-WALHours`

Ejemplo: **`EMRWALCount`**



Ejemplo**`EMRWALWorkspaceCount`**:

Etiquetado de espacios de trabajo de WAL

Puede añadir etiquetas a un espacio de trabajo al crear un nuevo espacio de trabajo y puede añadir, eliminar o enumerar las etiquetas de un espacio de trabajo activo para un clúster en ejecución. No puede etiquetar los recursos individuales del espacio de trabajo ni actualizar las etiquetas existentes; en su lugar, elimine las etiquetas no deseadas del espacio de trabajo y sustitúyalas.

Puede etiquetar los espacios de trabajo desde EMRWAL CLI. Para obtener una lista de comandos de EMRWAL CLI para etiquetar espacios de trabajo, consulte Referencia de la (EMRWAL) CLI de Amazon EMR WAL.

El siguiente ejemplo de política de IAM ilustra un escenario que permite realizar operaciones CRUDL en el espacio de trabajo únicamente con la clave `resource_tag_allow_test_key` y el valor `resource_tag_allow_test_value` de etiquetado adecuados:

```
{
  "Version": "2012-10-17",
  "Statement": [
      {
          "Action": [
              "EMRWAL:*"
          ],
          "Effect": "Allow",
          "Resource": [
              "*",
              "*"
          ],
          "Condition": {
              "StringEquals": {
                  "aws:ResourceTag/resource_tag_allow_test_key": [
                      "resource_tag_allow_test_value"
                  ]
              }
          }
      }
  ]
}
```

Para comprobar que la etiqueta ahora es necesaria para las operaciones del espacio de trabajo, utilice Referencia de la (EMRWAL) CLI de Amazon EMR WAL para llamar al comando `listTagsForResource` en `tagAllowResourceTag` para activar el espacio de trabajo con la

etiqueta de recurso deseada. Si configuró la condición correctamente, el comando se ejecutará correctamente.

```
emrwal listTagsForResource -r us-east-1 -arn arn:aws:emrwal:us-east-1:arn:workspace/
tagAllowResourceTag
Tag(Key=resource_tag_allow_test_key, Value=resource_tag_allow_test_value)
```

Replicación entre clústeres EMR WAL

A partir de EMR 7.5, EMR WAL admite la HBase replicación entre clústeres de registros de escritura anticipada. En este tema, se muestra cómo habilitar la función y comprobar que funciona. Para obtener más información sobre la replicación de clústeres, consulte la replicación de clústeres en la documentación de Apache. HBase

> **Note**
>
> Los registros de escritura anticipada conllevan un costo de lectura adicional, ya que el proceso de replicación lee los datos del EMR WAL local. Para obtener más información sobre el costo, consulte Acerca de las versiones de Amazon EMR.

Configuración de replicación entre clústeres

Para tener la misma experiencia de usuario, habilitar la función de replicación en EMR WAL es lo mismo que con los registros nativos de HBase escritura anticipada. El procedimiento que sigue muestra un ejemplo sencillo. Para obtener más información, consulte la HBase documentación en Cluster Replication.

1. Inicie un clúster principal, que es la fuente de replicación, con EMR WAL habilitado. Para habilitar los registros de escritura anticipada, consulte. Habilitación de Amazon EMR WAL Además, inicie un clúster homólogo. Para este clúster homólogo, puede optar por habilitar EMR WAL o no.
2. En ambos clústeres, cree una tabla:

   ```
   HBASE_CMD="sudo -u hbase hbase"
   echo "create 'test_replication_table',{NAME => 'CF'}" | $HBASE_CMD shell
   ```

3. Agregue una configuración de pares en el clúster principal y habilite la replicación de la tabla. Durante la adición de pares, necesita el nombre de host del nodo principal del clúster homólogo, que es PEER_DNS.

```
HBASE_CMD="sudo -u hbase hbase"
PEER_DNS="ip-10-1-1-0.ec2.com"
PEER_NAME="aws"
TABLE_NAME="test_replication_table"

## Create peering with the destination cluster
echo "add_peer '$PEER_NAME', CLUSTER_KEY => '$PEER_DNS:2181:/hbase'" | $HBASE_CMD
 shell

## List peers in the primary cluster to confirm peer setup
echo "list_peers" | $HBASE_CMD shell

## Enable table replication
echo "enable_table_replication '$TABLE_NAME'" | $HBASE_CMD shell
```

Confirmar la replicación entre clústeres

Tras realizar los pasos de configuración, se habilita la replicación entre el clúster principal y el clúster homólogo. A continuación, se realiza una prueba que confirma que la replicación funciona.

1. Agregue datos en el clúster principal y verifique los datos replicados en el clúster homólogo.

```
## Write on primary cluster with HBase CLI

put 'test_replication_table', 'aaa', 'CF:a', 'aaa_a1'
put 'test_replication_table', 'bbb', 'CF:b', 'bbb_b1'
put 'test_replication_table', 'ccc', 'CF:c', 'ccc_c1'
```

2. Confirme que la replicación en el clúster homólogo se ha realizado correctamente. En este caso, debería ver los datos replicados escritos desde el clúster principal al clúster homólogo.

```
### Scan on peer cluster with HBase CLI

scan 'test_replication_table'
```

Consideraciones y regiones para Amazon EMR WAL

Consideraciones para Amazon EMR WAL

En la siguiente lista se describen las consideraciones y limitaciones importantes de Amazon EMR WAL:

- Amazon EMR WAL está disponible para su uso con las versiones 6.15.0 y posteriores de Amazon EMR.
- Amazon EMR WAL es un servicio opcional de pago. Usted paga por lo que usa: lectura, escritura y almacenamiento de datos. Para obtener más información, consulte Descripción de los precios y las métricas de Amazon EMR WAL y la página Precios de Amazon EMR.
- Amazon EMR WAL utiliza HBase Write Ahead Log (WAL). Para usar Amazon EMR WAL, sus clústeres deben usar HBase WAL.
- Para habilitar Amazon EMR WAL al crear un clúster, debe contar con los permisos de rol necesarios. Para obtener más información, consulte Uso de roles vinculados a servicios para el registro anticipado.
- Debe habilitar Amazon EMR WAL al crear el clúster con la, o la API AWS Management Console AWS CLI, y debe usar la configuración de grupos de instancias. No puede habilitar Amazon EMR WAL en un clúster en ejecución si no creó el clúster con Amazon EMR WAL. Tampoco puede editar las configuraciones `hbase-site` para habilitar Amazon EMR WAL en un clúster en ejecución.
- Solo puede habilitar Amazon EMR WAL en clústeres que usen Amazon S3 como directorio raíz.
- Antes de la versión 7.5.0 de Amazon EMR, los registros en Amazon EMR WAL tenían que ocupar 4 MB o menos. Sin embargo, con Amazon EMR versión 7.5.0 y versiones posteriores, el tamaño máximo del registro en EMR WAL se puede configurar mediante la propiedad. `emr.wal.max.payload.size` El valor predeterminado es 1 GB. El siguiente ejemplo establece el tamaño máximo del registro en 2 GB:

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "emr.wal.max.payload.size": "2147483648"
    }
  }
]
```

- No puede tener varios clústeres activos en el mismo directorio HBase raíz de Amazon S3.
- No puede habilitar Amazon EMR WAL en clústeres de réplicas de lectura.
- El WAL se replica en todas las zonas de disponibilidad del servicio administrado.
- WAL perdura más allá del clúster y permanece disponible para el siguiente clúster..
- No puede deshabilitar Amazon EMR WAL durante el lanzamiento o cuando el clúster esté operativo (en estado de ejecución).
- Para obtener información sobre WAL y los límites del espacio de trabajo, consulte puntos de conexión y cuotas de Amazon EMR.

Disponibilidad regional para Amazon EMR WAL

El servicio Amazon EMR WAL está disponible en los siguientes lugares: Regiones de AWS

- `ap-northeast-1`: Asia Pacífico (Tokio)
- `ap-northeast-2`: Asia-Pacífico (Seúl)
- `ap-southeast-1`: Asia Pacífico (Singapur)
- `ap-south-1`: Asia-Pacífico (Bombay)
- `ap-southeast-2`: Asia Pacífico (Sídney)
- `eu-central-1`: Europa (Fráncfort)
- `eu-north-1`: Europa (Estocolmo)
- `eu-west-1`: Europa (Irlanda)
- `sa-east-1`: América del Sur (São Paulo)
- `us-east-1`: Este de EE. UU. (Norte de Virginia)
- `us-east-2`: Este de EE. UU. (Ohio)
- `us-west-2`: Oeste de EE. UU. (Oregón)

Las siguientes regiones solo estarían disponibles con Amazon EMR versión 7.3.0 o posterior:

- `ap-east-1`: Asia-Pacífico (Hong Kong)
- `af-south-1`: África (Ciudad del Cabo)
- `ca-central-1`: Canadá (centro)

- `eu-west-2`: Europa (Londres)

Referencia de la (EMRWAL) CLI de Amazon EMR WAL

La interfaz de línea de comandos (EMRWAL CLI)EMRWAL es una herramienta unificada para administrar el registro de escritura anticipada (WAL) para Amazon EMR. La EMRWAL CLI se envía con los clústeres de EMR cuando se habilita WAL en el momento de crear un clúster. Para obtener más información acerca de la habilitación de WAL, consulte Write Ahead Logs (WAL) para Amazon EMR.

La EMRWAL CLI incluye los siguientes comandos:

Temas

- createWorkspace
- deleteWal
- deleteWorkspace
- listTagsForResource
- listWals
- listWorkspaces
- tagResource
- untagResource

`createWorkspace`

El comando `createWorkspace` crea un nuevo espacio de trabajo de Amazon EMR WAL.

Uso:

```
emrwal createWorkspace [-tags <tags>] [-e {endpoint}] [-r {Region}] -w {workspacename}
 [-h]
```

Ejemplo:

```
emrwal createWorkspace -w examplews
```

deleteWal

El comando `deleteWals` elimina el Amazon EMR WAL que especifique.

Uso:

```
emrwal deleteWal [-e {endpoint}] [-r {Region}] [-w {workspacename}] [-p <tablePrefix>]
 [-n <walName>] [-N <fullName>] [-R] [-m] [-h]
```

Ejemplo:

```
emrwal deleteWal -w examplews -p hbasetable -n examplewal
```

deleteWorkspace

El comando `deleteWorkspace` elimina el espacio de trabajo de Amazon EMR WAL que especifique.

Uso:

```
emrwal deleteWorkspace [-e {endpoint}] [-r {Region}] -w {workspacename} [-h]
```

Ejemplo:

```
emrwal deleteWorkspace -w examplews
```

listTagsForResource

El comando `listTagsForResource` enumera todas las etiquetas de pares clave-valor del espacio de trabajo de Amazon EMR WAL que especifique.

Uso:

```
emrwal listTagsForResource -arn {resource-arn} [-e {endpoint}] [-r {Region}] [-h]
```

Ejemplo:

```
emrwal listTagsForResource -arn arn:aws:emrwal::1234567891234:workspace/examplews
```

`listWals`

El `listWals` comando muestra todos los Amazon EMR del espacio WALs de trabajo que especifique.

Uso:

```
emrwal listWals [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}] [-
r {Region}] [-w {workspacename}] [-p <tablePrefix>] [-M {integer}] [-h]
```

Ejemplo:

```
emrwal listWals -w examplews
```

`listWorkspaces`

El comando `listWorkspaces` enumera todos los espacios de Amazon EMR WAL que están disponibles para usted.

Uso:

```
emrwal listWorkspaces [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}]
 [-r {Region}] [-M {integer}] [-h]
```

Ejemplo:

```
emrwal listWorkspaces
```

`tagResource`

El comando `tagResource` asigna una o más etiquetas de pares clave-valor al espacio de trabajo de Amazon EMR WAL que especifique.

Uso:

```
emrwal tagResource -arn {resource-arn} -tags <tags> [-e {endpoint}] [-r {Region}] [-h]
```

Ejemplo:

```
emrwal tagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tags tag_key=tag_value
```

`untagResource`

El comando `untagResource` anula la asignación de una o más etiquetas de pares clave-valor al espacio de trabajo Amazon EMR WAL que especifique.

Uso:

```
emrwal untagResource -arn {resource-arn} -tagKeys <tagKeys> [-e {endpoint}] [-
r {Region}] [-h]
```

Ejemplo:

```
emrwal untagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tagKeys tag_key
```

Uso del HBase shell

Después de crear un HBase clúster, el siguiente paso es conectarse para HBase poder empezar a leer y escribir datos (los clústeres de lectura y réplica no admiten la escritura de datos). Puede usar el HBase shell para probar los comandos.

Para abrir la HBase carcasa

1. Use SSH para conectarse al servidor principal del HBase clúster. Para obtener más información sobre cómo conectarse al nodo principal de Amazon EMR mediante SSH, consulte Conectarse al nodo principal de Amazon EMR mediante SSH en la Guía de administración de Amazon EMR.
2. Ejecute `hbase shell`. El HBase shell se abre con un mensaje similar al siguiente.

   ```
   hbase(main):001:0>
   ```

Puede ejecutar comandos de HBase shell desde la línea de comandos. Para obtener más información sobre los comandos del shell y cómo llamarlos, escriba help en la HBase línea de comandos y pulse Entrar.

Creación de una tabla

El siguiente comando crea una tabla denominada “t1” que tiene una única familia de columnas denominada “f1”.

```
hbase(main):001:0>create 't1', 'f1'
```

Colocar un valor

El siguiente comando coloca el valor “v1” para la fila “r1” en la tabla “t1” y la columna “f1”.

```
hbase(main):001:0>put 't1', 'r1', 'f1:col1', 'v1'
```

Obtener un valor

El siguiente comando obtiene los valores para la fila “r1” de la tabla “t1”.

```
hbase(main):001:0>get 't1', 'r1'
```

Eliminación de una tabla

El siguiente comando elimina y borra la tabla “t1”.

```
hbase(main):001:0>drop 'ns1:t1',false
```

El valor booleano corresponde a si desea archivar la tabla o no, por lo que puede configurarlo como `true` si desea guardarla. También puede ejecutar `drop 'ns1:t1'` sin ningún booleano para archivar la tabla.

Acceda a HBase las tablas con Hive

HBase y Apache Hive están perfectamente integrados, lo que le permite ejecutar cargas de trabajo de procesamiento masivo en paralelo directamente sobre los datos almacenados en. HBase Para usar Hive con ellas HBase, normalmente puedes lanzarlas en el mismo clúster. Sin embargo, puedes lanzar Hive y HBase en clústeres independientes. Ejecutar HBase Hive por separado en distintos clústeres puede mejorar el rendimiento, ya que permite que cada aplicación utilice los recursos del clúster de forma más eficiente.

Los siguientes procedimientos muestran cómo conectarse a HBase un clúster mediante Hive.

> **Note**
>
> Solo puede conectar un clúster de Hive a un único HBase clúster.

Para conectar Hive a HBase

1. Cree clústeres independientes con Hive e HBase instalado o cree un solo clúster con ambos HBase y Hive instalado.
2. Si utiliza clústeres independientes, modifique los grupos de seguridad para que HBase los puertos de Hive estén abiertos entre estos dos nodos principales.
3. Utilice SSH para conectar al nodo principal para el clúster con Hive instalado. Para obtener más información, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR.
4. Lance el shell Hive con el siguiente comando.

   ```
   hive
   ```

5. (Opcional) No es necesario que lo haga si HBase Hive se encuentra en el mismo clúster. Conecte el HBase cliente de su clúster de Hive al HBase clúster que contiene sus datos. En el siguiente ejemplo, *public-DNS-name* se reemplaza por el nombre DNS público del nodo principal del HBase clúster, por ejemplo:`ec2-50-19-76-67.compute-1.amazonaws.com`.

   ```
   set hbase.zookeeper.quorum=public-DNS-name;
   ```

6. Proceda a ejecutar las consultas de Hive en sus HBase datos según lo desee o consulte el siguiente procedimiento.

Para acceder a HBase los datos de Hive

- Una vez realizada la conexión entre el Hive y HBase los clústeres (como se muestra en el procedimiento anterior), puede acceder a los datos almacenados en el HBase clúster creando una tabla externa en Hive.

El siguiente ejemplo, cuando se ejecuta desde la línea de comandos de Hive en el nodo principal, crea una tabla externa que hace referencia a los datos almacenados en una HBase tabla llamada. `inputTable` A continuación, puede hacer referencia `inputTable` a las instrucciones de Hive para consultar y modificar los datos almacenados en el HBase clúster.

```
set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com;

create external table inputTable (key string, value string)
     stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
      with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
      tblproperties ("hbase.table.name" = "t1");

select count(key) from inputTable ;
```

Para ver un caso de uso más avanzado y un ejemplo de combinación HBase y Hive, consulte la entrada del blog sobre AWS big data, Combine NoSQL y massively parallel analytics mediante HBase Apache y Apache Hive en Amazon EMR.

Uso de instantáneas HBase

HBase utiliza una función de instantáneas integrada para crear copias de seguridad ligeras de las tablas. En los clústeres de EMR, estas copias de seguridad se pueden exportar a Amazon S3 mediante EMRFS. Puede crear una instantánea en el nodo principal mediante el HBase shell. En este tema se muestra cómo ejecutar estos comandos de forma interactiva con el shell o mediante un paso mediante `command-runner.jar` el comando AWS CLI o AWS SDK para Java. Para obtener más información sobre otros tipos de HBase copias de seguridad, consulte la sección de copias HBase de seguridad en la HBase documentación.

Crear una instantánea con una tabla

```
hbase snapshot create -n snapshotName -t tableName
```

Uso de command-runner.jar desde AWS CLI:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
```

```
Args=[ "hbase", "snapshot", "create","-n","snapshotName","-t","tableName"]
```

AWS SDK para Java

```
HadoopJarStepConfig hbaseSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","create","-n","snapshotName","-t","tableName");
```

> **Note**
>
> Si su nombre de instantánea no es único, la operación de creación devuelve un código `-1` o `255`, pero es posible que no vea un mensaje de error que indique que se realizó incorrectamente. Para utilizar el mismo nombre de instantánea, elimínelo y, a continuación, vuelva a crearlo.

Eliminar una instantánea

```
hbase shell
>> delete_snapshot 'snapshotName'
```

Ver información de instantánea

```
hbase snapshot info -snapshot snapshotName
```

Exportar una instantánea a Amazon S3

> **Important**
>
> Si no especifica un `-mappers` valor al exportar una instantánea, HBase utiliza un cálculo arbitrario para determinar el número de mapeadores. Este valor pueden ser muy grande en función del tamaño de la tabla, lo que afecta negativamente a la ejecución de trabajos durante la exportación. Por este motivo, le recomendamos que especifique el parámetro `-mappers`, el parámetro `-bandwidth` (que especifica el consumo de ancho de banda en megabytes por segundo) o ambos para limitar los recursos del clúster utilizados por la operación de exportación. De forma alternativa, puede ejecutar la operación de exportación de instantánea durante un periodo de bajo consumo.

```
hbase snapshot export -snapshot snapshotName \
-copy-to s3://amzn-s3-demo-bucket/folder -mappers 2
```

Se utiliza `command-runner.jar` desde: AWS CLI

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "export","-snapshot","snapshotName","-copy-to","s3://amzn-s3-demo-bucket/folder","-mappers","2","-bandwidth","50"]
```

AWS SDK para Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","export",
      "-snapshot","snapshotName","-copy-to",
      "s3://bucketName/folder",
      "-mappers","2","-bandwidth","50");
```

Importar instantáneas desde Amazon S3

Aunque se trata de una importación, la HBase opción utilizada aquí sigue siéndoloexport.

```
sudo -u hbase hbase snapshot export \
-D hbase.rootdir=s3://amzn-s3-demo-bucket/folder \
-snapshot snapshotName \
-copy-to hdfs://masterPublicDNSName:8020/user/hbase \
-mappers 2
```

Usando `command-runner.jar` desde AWS CLI:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar", \
Args=["sudo","-u","hbase","hbase snapshot export","-snapshot","snapshotName", \
"-D","hbase.rootdir=s3://amzn-s3-demo-bucket/folder", \
"-copy-to","hdfs://masterPublicDNSName:8020/user/hbase","-mappers","2","-chmod","700"]
```

AWS SDK para Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
```

```
  .withJar("command-runner.jar")
  .withArgs("sudo","-u","hbase","hbase","snapshot","export", "-D","hbase.rootdir=s3://
path/to/snapshot",
      "-snapshot","snapshotName","-copy-to",
      "hdfs://masterPublicDNSName:8020/user/hbase",
      "-mappers","2","-chuser","hbase");
```

Restaurar una tabla a partir de instantáneas del shell HBase

```
hbase shell
>> disable tableName
>> restore_snapshot snapshotName
>> enable tableName
```

HBase actualmente no admite todos los comandos de instantáneas que se encuentran en el HBase shell. Por ejemplo, no existe una opción de HBase línea de comandos para restaurar una instantánea, por lo que debe restaurarla dentro de un shell. Esto significa que `command-runner.jar` debe ejecutar un comando Bash.

> **Note**
>
> Dado que el comando que se utiliza aquí es `echo`, es posible que el comando del intérprete de comandos siga devolviendo errores incluso si el comando ejecutado por Amazon EMR devuelve un código de salida `0`. Consulte los registros de pasos, si decide ejecutar un comando de shell como paso.

```
echo 'disable tableName; \
restore_snapshot snapshotName; \
enable tableName' | hbase shell
```

Este es el paso utilizando la AWS CLI. En primer lugar, cree el siguiente archivo `snapshot.json`:

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo $'disable \"tableName\"; restore_snapshot
 \"snapshotName\"; enable \"tableName\"' | hbase shell"],
    "Jar": "command-runner.jar",
```

```
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./snapshot.json
```

AWS SDK para Java:

```
HadoopJarStepConfig hbaseRestoreSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("bash","-c","echo $'disable \"tableName\"; restore_snapshot \"snapshotName
\"; enable \"snapshotName\"' | hbase shell");
```

Configuración HBase

Si bien la HBase configuración predeterminada debería funcionar para la mayoría de las aplicaciones, puede modificarla. HBase Para ello, utilice las propiedades de las clasificaciones de HBase configuración. Para obtener más información, consulte Configuración de aplicaciones.

El siguiente ejemplo crea un clúster con un directorio HBase raíz alternativo basado en un archivo de configuración`myConfig.json`, almacenado en Amazon S3.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=HBase \
--instance-type m5.xlarge --instance-count 3 --configurations https://s3.amazonaws.com/
amzn-s3-demo-bucket/myfolder/myConfig.json
```

El archivo `myConfig.json` especifica la propiedad `hbase.rootdir` de la clasificación de configuración `hbase-site`, tal y como se muestra en el siguiente ejemplo. *`ip-XXX-XX-XX-XXX.ec2.internal`*Sustitúyalo por el nombre de host DNS interno del nodo principal del clúster.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.rootdir": "hdfs://ip-XXX-XX-XX-XXX.ec2.internal:8020/user/
myCustomHBaseDir"
    }
  }
]
```

> **Note**
>
> Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Cambios en la asignación de memoria en YARN

HBase no se ejecuta como una aplicación de YARN, por lo que es necesario volver a calcular la memoria asignada a YARN y sus aplicaciones, lo que se traduce en una reducción de la memoria total disponible para YARN si HBase está instalada. Debe tener esto en cuenta a la hora de planificar la ubicación conjunta de las aplicaciones de YARN y HBase en los mismos clústeres. Los tipos de instancias con menos de 64 GB de memoria tienen la mitad de la memoria disponibleNodeManager, que luego se asigna a. HBase RegionServer En el caso de los tipos con una memoria superior a 64 GB, HBase RegionServer la memoria está limitada a 32 GB. Como regla general, la memoria de configuración de YARN es un múltiplo de la memoria de tareas MapReduce reductora.

Las tablas de abajo Valores predeterminados para los ajustes de configuración de tareas muestran los cambios en la configuración de YARN en función de la memoria necesaria HBase.

HBase números de puerto

Algunos números de puerto elegidos HBase son diferentes de los predeterminados. A continuación se muestran las interfaces y los puertos para HBase Amazon EMR.

HBase puertos

Interfaz	Puerto	Protocolo
HMaster	16 000	TCP
HMaster INTERFAZ DE USUARIO	16010	HTTP
RegionServer	16020	TCP
RegionServer Información	16030	HTTP
Servidor REST	8070	HTTP
IU REST	8085	HTTP
Servidor Thrift	9090	TCP
IU Servidor Thrift	9095	HTTP

> **Important**
>
> `kms-http-port` es 9700 y `kms-admin-port` es 9701 en la versión de lanzamiento de Amazon EMR 4.6.0 y posteriores.

HBase configuración del sitio para optimizar

Puede establecer una o todas las configuraciones del HBase sitio para optimizar el HBase clúster para la carga de trabajo de la aplicación. Le recomendamos los siguientes ajustes como punto de partida en la investigación.

zookeeper.session.timeout

El tiempo de espera predeterminado es de 40 segundos (40 000 ms). Si un servidor de región se bloquea, esto es lo que tarda el servidor principal en advertir la ausencia de un servidor de región y comenzar la recuperación. Para ayudar a que el servidor principal se recupere con mayor rapidez, puede reducir este valor a un periodo de tiempo más corto. En el ejemplo siguiente se utilizan 30 segundos o 30 000 ms:

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "zookeeper.session.timeout": "30000"
    }
  }
]
```

hbase.regionserver.handler.count

Esto define el número de procesos que el servidor de región mantiene abiertos para servir las solicitudes a las tablas. El valor predeterminado de 10 es bajo, a fin de impedir que los usuarios finalicen sus servidores de región al utilizar grandes búferes de escritura con un elevado número de clientes simultáneos. La regla general es mantener este número bajo cuando la carga útil por solicitud se aproxime al rango de MB (grandes operaciones de venta, escaneos con una memoria caché grande) y alto cuando la carga útil es pequeña (opciones de venta pequeñas ICVs, eliminaciones). El siguiente ejemplo aumenta el número de subprocesos abiertos a 30:

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.regionserver.handler.count": "30"
    }
  }
]
```

hbase.hregion.max.filesize

Este parámetro controla el tamaño, en bytes, de las regiones individuales. De forma predeterminada, está establecido en `1073741824`. Si escribes muchos datos en tu HBase clúster y eso provoca que se divida con frecuencia, puedes aumentar este tamaño para agrandar las regiones individuales. Esto reduce las divisiones, pero se tarda más tiempo en equilibrar la carga de las regiones de un servidor a otro.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
```

```
       "hbase.hregion.max.filesize": "1073741824"
    }
  }
]
```

hbase.hregion.memstore.flush.size

Este parámetro controla el tamaño máximo de memstore, en bytes, antes de que se vacíe al disco. De forma predeterminada, es 134217728. Si la carga de trabajo consta de ráfagas cortas de operaciones de escritura, es posible que desee aumentar este límite para que todas las escrituras se mantengan en la memoria durante la ráfaga y se vacíen al disco después. Esto puede mejorar el rendimiento durante las ráfagas.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.hregion.memstore.flush.size": "134217728"
    }
  }
]
```

Consideraciones sobre el rendimiento

Habilitar un recolector de basura (ZGC) para HBase

Con Amazon EMR versión 7.10.0 y versiones posteriores, los usuarios pueden configurar sin problemas los ajustes de recolección de basura (GC) para su clúster. HBase Recomendamos habilitar Z Garbage Collector (ZGC) para lograr tiempos de pausa de GC de baja latencia y de menos de milisegundos.

Esta es una configuración para habilitar el ZGC para: HBase RegionServer

```
[
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
```

```
                "JAVA_HOME": "/usr/lib/jvm/jre-21",
                "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational\""
            }
        }
    ]
  }
]
```

> **Note**
>
> La variable de entorno JAVA_HOME debe configurarse cuando se utiliza el ZGC generacional (recomendado), ya que se introdujo en el JDK 21. Si desea utilizar el modo no generacional (sin él**`-XX:+ZGenerational`**) del ZGC, no es necesario configurar el JAVA_HOME. En el JDK 24, se ha eliminado el modo no generacional del ZGC.

Ajuste ZGC

1. Habilite el montón fijo de JVM

   El recolector de basura Z (ZGC) funciona de manera más eficiente cuando se configura con una memoria de pila fija, lo que elimina la sobrecarga que supone devolver la memoria al sistema operativo. Para configurar la memoria de pila fija para el HBase clúster, utilice la siguiente configuración:

   ```
   [
     {
       "Classification": "hbase",
       "Properties": {
         "hbase.regionserver.fixed.heap.enabled": "true"
       }
     }
   ]
   ```

2. Habilite la función Pretouch

   La activación del pretoque mejora el rendimiento de la recolección de basura (GC) al reducir la latencia inicial y ofrecer un rendimiento más predecible. Para activar el pretoque en el HBase clúster, utilice la siguiente configuración:

   ```
   [
   ```

```
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
                    "JAVA_HOME": "/usr/lib/jvm/jre-21",
                    "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational -
XX:+AlwaysPreTouch\""
                }
            }
        ]
    }
]
```

Vea la interfaz de HBase usuario

> **Note**
>
> La interfaz HBase de usuario utiliza conexiones HTTP inseguras de forma predeterminada. Para habilitar el HTTP seguro (HTTPS), defina la `hbase.ssl.enabled` propiedad de la `hbase-site` clasificación como `true` en su HBase configuración. Para obtener más información sobre el uso de HTTP (HTTPS) seguro para la interfaz de usuario HBase web, consulte la Guía de HBase referencia de Apache.

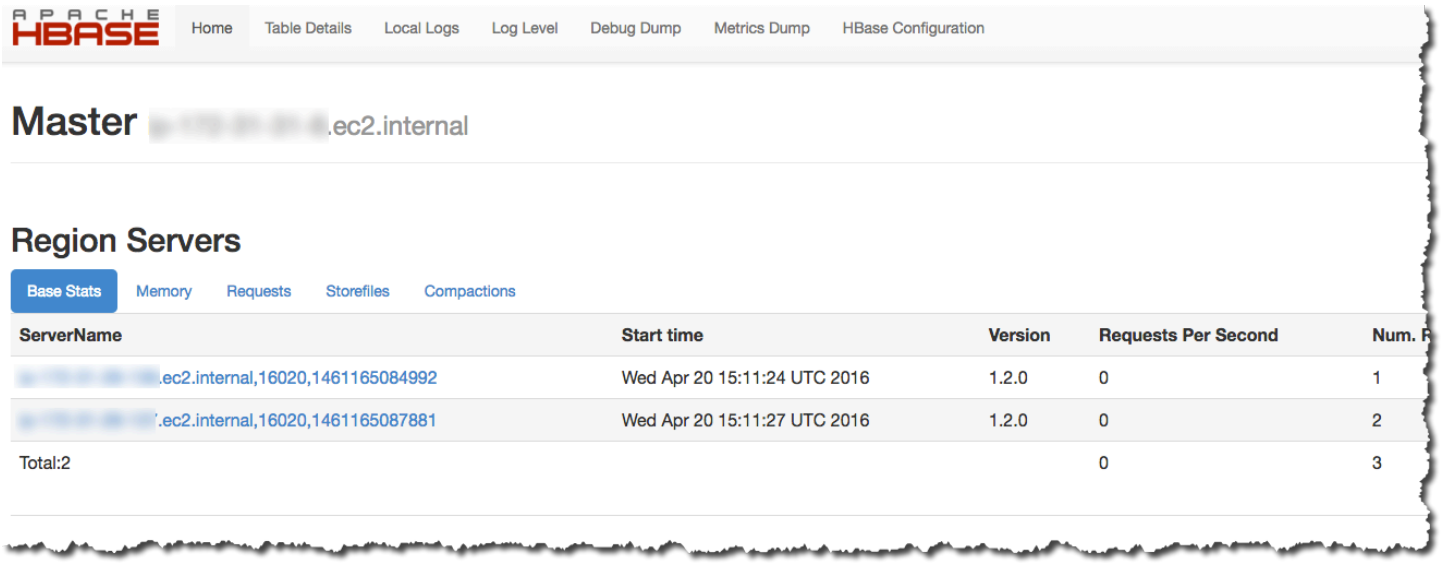
HBase proporciona una interfaz de usuario basada en la web que puede utilizar para supervisar el HBase clúster. Cuando se ejecuta HBase en Amazon EMR, la interfaz web se ejecuta en el nodo principal y se puede ver mediante el reenvío de puertos, también conocido como creación de un túnel SSH.

Para ver la interfaz de usuario HBase

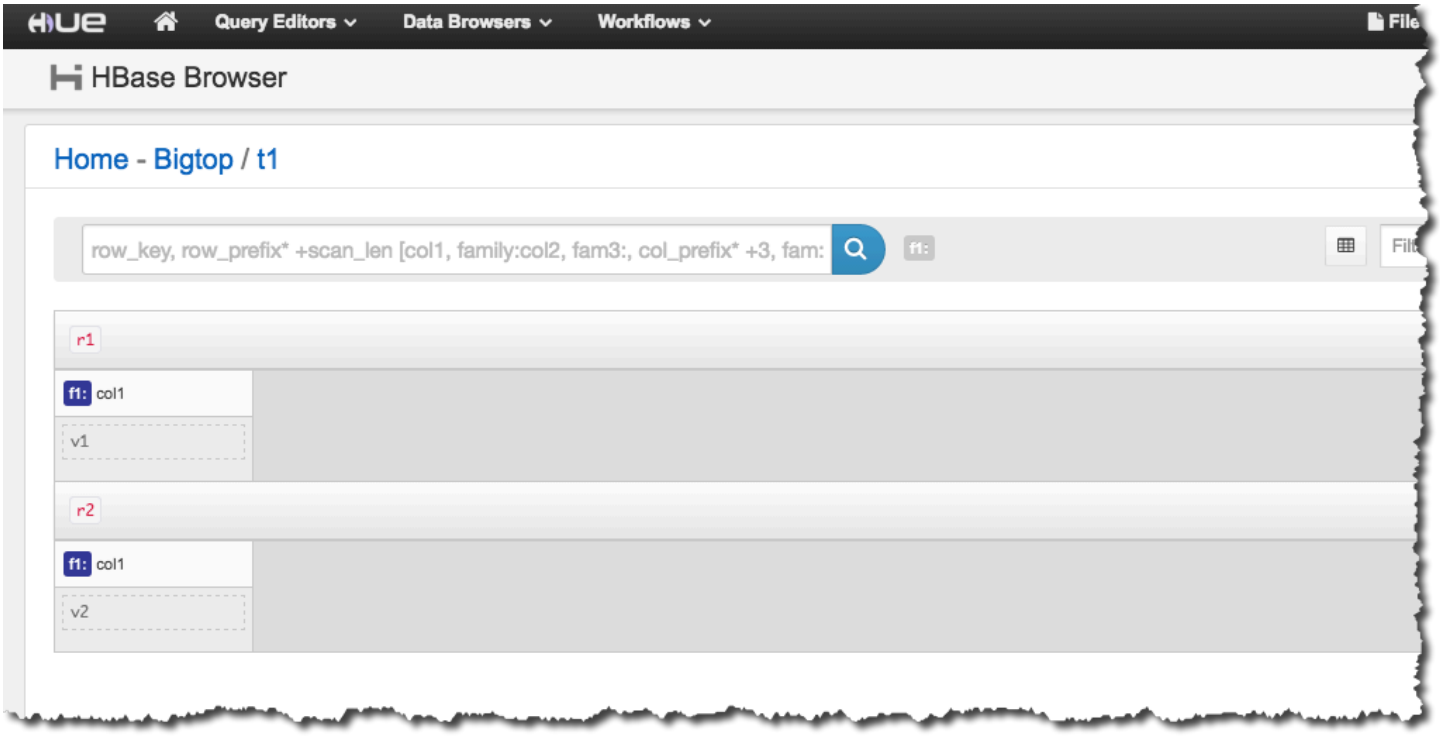
1. Utilice SSH para crear un túnel en el nodo principal y crear una conexión segura. Para obtener más información, consulte Opción 2, parte 1: configurar un túnel SSH hacia el nodo maestro mediante el reenvío dinámico de puertos en la Guía de administración de Amazon EMR.
2. Instale un navegador web con una herramienta de proxy, como el FoxyProxy complemento para Firefox, a fin de crear un proxy SOCKS para AWS los dominios. Para obtener más información,

consulte Opción 2, parte 2: configurar los ajustes del proxy para ver los sitios web alojados en el nodo maestro en la Guía de administración de Amazon EMR.

3. Con el proxy configurado y la conexión SSH abierta, puede ver la HBase interfaz de usuario abriendo una ventana del navegador con http: //:16010/master-status***master-public-dns-name***, donde se encuentra *master-public-dns-name* la dirección DNS pública del nodo principal del clúster.



También puedes verla en Hue. HBase Por ejemplo, lo siguiente muestra la tabla, `t1`, creada en Uso del HBase shell:



Para obtener más información Hue, consulte Hue.

Ver archivos de HBase registro

Como parte de su funcionamiento, HBase escribe archivos de registro con detalles sobre los ajustes de configuración, las acciones de los demonios y las excepciones. Estos archivos de registro pueden ser útiles para depurar problemas HBase y para realizar un seguimiento del rendimiento.

Si configura el clúster para mantener los archivos de registro en Amazon S3, debe saber que los registros se escriben en Amazon S3 cada cinco minutos, por lo que es posible que haya un pequeño retraso antes de que los últimos archivos de registro estén disponibles.

Para ver HBase los registros del nodo principal

- Puede ver los HBase registros actuales mediante SSH para conectarse al nodo principal y navegar hasta el `/var/log/hbase` directorio. Estos registros no estarán disponibles una vez que el clúster se haya terminado a menos que hubiera habilitado el registro en Amazon S3 cuando se lanzó el clúster.

Para ver HBase los registros en Amazon S3

- Para acceder a HBase los registros y otros registros de clústeres en Amazon S3 y tenerlos disponibles una vez finalizado el clúster, especifique un bucket de Amazon S3 para recibir estos registros al crear el clúster. Esto se realiza mediante la opción `--log-uri`. Para obtener más información sobre la habilitación del registro para el clúster, consulte Configurar el registro de clúster y depuración (opcional) en la Guía de administración de Amazon EMR.

Supervise HBase con Ganglia

> **Note**
>
> La última versión de Amazon EMR que incluye Ganglia es Amazon EMR 6.15.0. Para monitorizar el clúster, las versiones posteriores a la 6.15.0 incluyen el agente de Amazon CloudWatch .

El proyecto de código abierto Ganglia es un sistema distribuido y escalable diseñado para monitorizar clústeres y redes al mismo tiempo que minimiza el impacto en su rendimiento. Al habilitar Ganglia en su clúster, puede generar informes y ver el rendimiento del clúster en su conjunto, además de inspeccionar el rendimiento de cada una de las instancias de nodo individuales. Para obtener más información acerca del proyecto de código abierto Ganglia, consulte http://ganglia.info/. Para obtener más información sobre el uso de Ganglia con clústeres de Amazon EMR, consulte Ganglia.

Una vez que el clúster se haya lanzado con Ganglia configurado, puede acceder a los informes y gráficos de Ganglia utilizando la interfaz gráfica que se ejecuta en el nodo principal.

Ganglia almacena los archivos de log en el nodo principal del directorio `/mnt/var/lib/ganglia/rrds/`. En versiones anteriores de Amazon EMR, puede almacenar los archivos de registro en el directorio `/var/log/ganglia/rrds/`.

Para configurar un clúster para Ganglia y usar el HBase AWS CLI

- Use un comando `create-cluster` similar al siguiente:

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase Name=Ganglia --use-default-roles \
```

```
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3
```

> **Note**
>
> Si el rol de servicio Amazon EMR predeterminado y el perfil de EC2 instancia de Amazon no existen, se produce un error. Utilice el comando `aws emr create-default-roles` para crearlos y, a continuación, vuelva a intentarlo.

Para más información, consulte Comandos de Amazon EMR en la AWS CLI.

Para ver HBase las métricas en la interfaz web de Ganglia

1. Utilice SSH para crear un túnel en el nodo principal y crear una conexión segura. Para obtener más información, consulte Opción 2, parte 1: configurar un túnel SSH hacia el nodo maestro mediante el reenvío dinámico de puertos en la Guía de administración de Amazon EMR.
2. Instale un navegador web con una herramienta de proxy, como el FoxyProxy complemento para Firefox, para crear un proxy SOCKS para AWS los dominios. Para obtener más información, consulte Opción 2, parte 2: configurar los ajustes del proxy para ver los sitios web alojados en el nodo maestro en la Guía de administración de Amazon EMR.
3. Con el proxy configurado y la conexión SSH abierta, puede ver las métricas de Ganglia abriendo una ventana del navegador con http://*master-public-dns-name*/ganglia/, donde *master-public-dns-name* se encuentra la dirección DNS pública del servidor maestro del clúster. HBase

Para ver archivos de registro de Ganglia en el nodo principal

- Si el clúster se sigue ejecutando, puede acceder a los archivos de registro utilizando SSH para conectarse al nodo principal y navegando al directorio `/mnt/var/lib/ganglia/rrds/`. En EMR 3.x, vaya al directorio `/var/log/ganglia/rrds`. Para obtener más información, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR.

Para consultar los archivos de registro de Ganglia en Amazon S3

- Los archivos de registro de Ganglia no se escriben automáticamente en Amazon S3, aunque habilite el registro para el clúster. Para ver los archivos de registro de Ganglia en Amazon S3, debe insertar manualmente los registros desde `/mnt/var/lib/ganglia/rrds/` en el bucket de S3.

Supervisión de EMR con Amazon HBase CloudWatch

A partir de EMR 7.0, Amazon EMR proporciona el Amazon CloudWatch Agent para enviar las métricas o CloudWatch Prometheus, en sustitución del sistema de monitorización Ganglia. Encontrarás más detalles en la Guía de configuración de Amazon CloudWatch Agent.

El CloudWatch agente Amazon EMR 7.0 proporcionaba una integración básica y su configuración requería un esfuerzo manual, como descargar componentes y utilizar acciones de arranque. Sin embargo, a partir de EMR 7.1, el proceso se simplifica. El uso de Amazon CloudWatch Agent en EMR 7.1 facilita la supervisión de las métricas específicas del sistema y de la aplicación, incluidas las de. HBase Al usar la API de configuración de EMR, puede configurar y personalizar rápidamente el proceso de recopilación de métricas y elegir dónde enviar los datos, ya sea a Amazon CloudWatch o Prometheus. Esta flexibilidad le ayuda a vigilar de cerca sus HBase clústeres, lo que garantiza que funcionen sin problemas y de manera eficiente.

Uso de la API de configuración de EMR para gestionar toda la configuración, lo que facilita mucho el proceso. El CloudWatch agente de EMR 7.1 admite tres tipos principales de métricas:

- Métricas del sistema: incluyen indicadores clave del rendimiento del sistema, como el uso de la CPU, el uso del disco, el uso de la memoria, las E/S de la red, los procesos y el uso del intercambio.
- Métricas de Hadoop Daemon: estas métricas se refieren a los distintos componentes de Hadoop, que incluyen: DataNode métricas, métricas, métricas YARN y NameNode métricas YARN NodeManager . ResourceManager
- HBase Métricas: estas métricas proporcionan información sobre el rendimiento: métricas HBase maestras, métricas HBase de servidores HBase regionales, métricas de servidores HBase REST y métricas de Thrift Server. HBase

Using the AWS CLI

Todas las métricas de Hadoop HBase están basadas en JMX, lo que significa que utilizan Java Management Extensions para proporcionar información detallada. Así es como puedes configurar Amazon CloudWatch Agent para que supervise HBase:

- Consulta los requisitos previos antes de crear un clúster con Amazon CloudWatch Agent. Utilice un `create-cluster` comando similar al ejemplo que aparece después de esta lista.
- Consulte las configuraciones compatibles con un HBase clúster.
- Consulte el ejemplo siguiente para ver ejemplos de configuraciones para configurar la HBase supervisión. Consulte los ejemplos de configuraciones de la `--configuration` entrada.

```
aws emr create-cluster --name "HBase cluster with CloudWatch agent" \
--release-label emr-7.1.0 \
--applications Name=HBase Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--configurations file://./configurations.json \
--instance-count 3 --use-default-roles
```

Para obtener más información sobre los ejemplos de configuraciones de JSON para las métricas, consulta Configurar métricas.

Using the console

Para crear un clúster con el CloudWatch agente de Amazon desde la consola, lleve a cabo estos pasos:

Cree un clúster con el CloudWatch agente desde la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Create cluster.
3. En Nombre y aplicaciones, elija la versión de Amazon EMR 7.0.0 o posterior.
4. En Paquete de aplicaciones, seleccione HBase las aplicaciones adicionales que desee instalar en su clúster e incluya el CloudWatch agente en sus selecciones.
5. Amplíe la configuración del software. A continuación, puede introducir la configuración directamente a través de JSON o una sintaxis abreviada mostrada en el texto en sombra en la consola. De lo contrario, puede proporcionar una URI de Amazon S3 para un archivo con

un objeto `Configurations` de JSON. Para obtener más información sobre los ejemplos de configuraciones de métricas en JSON, consulte Configurar métricas.

6. Continúe con la creación del clúster que se ajuste a las necesidades del caso de uso.

Configura las métricas

Para supervisar el HBase Master, puedes configurar Amazon CloudWatch Agent para que recopile métricas específicas.

1. Configuración de métricas HBase maestras: para monitorear la métrica HBase maestra, puedes configurar Amazon CloudWatch Agent para que recopile métricas específicas. Este es un ejemplo de configuración para realizar un seguimiento de las actividades del administrador de tareas principal:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-master-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Master,sub=AssignmentManager":
 "AssignFailedCount,AssignSubmittedCount",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

En esta configuración:

- Especificamos el MBean (`Hadoop:service=HBase,name=Master,sub=AssignmentManager`) para recopilar métricas como `AssignFailedCount` y`AssignSubmittedCount`.

- Establecemos el intervalo para recopilar estas métricas cada 30 segundos (30 000 milisegundos).

2. Configuración HBase de las métricas de los servidores regionales: para supervisar los servidores HBase regionales, configure el CloudWatch agente de la siguiente manera:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-region-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=RegionServer,sub=IPC":
 "numActiveHandler,numActivePriorityHandler",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Esta configuración:

- Supervisa los controladores activos en el servidor de región (`numActiveHandler`,`numActivePriorityHandler`).
- Utiliza un intervalo de 30 segundos para la recopilación de métricas.

3. Configuración de las métricas del servidor HBase REST: para monitorear la interfaz HBase REST, puede usar la siguiente configuración:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-rest-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=REST": "successfulPut,successfulScanCount",
```

```
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

En este ejemplo, el CloudWatch agente recopila métricas sobre las operaciones PUT correctas y escanea los recuentos cada 30 segundos.

4. Configuración de las métricas de HBase Thrift Server: para supervisar el HBase Thrift Server, puede configurar las métricas con una configuración como la siguiente:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-thrift-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Thrift,sub=ThriftOne":
 "BatchGet_max,BatchGet_mean",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

Esta configuración registra los tiempos máximos y medios de las operaciones GET por lotes en el servidor Thrift.

Uso del destino de las métricas

Con EMR 7.1, tiene la opción de enviar los datos de sus métricas a Amazon o CloudWatch Amazon Managed Service for Prometheus. Esta opción le permite integrarse sin problemas con diferentes herramientas de monitorización, en función de sus necesidades.

Envío de métricas a Amazon CloudWatch

Para enviar las métricas a CloudWatch, usa esta configuración:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

Envío de métricas a Amazon Managed Service para Prometheus

Si prefieres usar Prometheus, establece el destino y proporciona la URL del punto final:

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "https://aps-workspaces.region.amazonaws.com/
workspaces/workspace_id/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

`region`Sustitúyala por tu AWS región y por tu `workspace_id` ID de espacio de trabajo de Prometheus. Esta configuración dirige sus HBase métricas a Prometheus mediante el punto final especificado.

Con la configuración anterior, puede ver las siguientes métricas en la pestaña Supervisión.

¿Estás migrando desde versiones anteriores HBase

Para migrar datos de una HBase versión anterior, consulte Actualización, número de HBase versión y compatibilidad en la Guía de HBase referencia de Apache. Es posible que deba prestar especial atención a los requisitos para actualizar desde versiones anteriores a la 1.0 de. HBase

Migración a Amazon EMR versión 7.4.0 o posterior

> **Note**
>
> Siga estas pautas si va a migrar de una versión EMR anterior a la 7.4.0 a una versión posterior a la 7.3.0.

Si actualmente utilizas una versión de EMR con la función de seguimiento de archivos de tienda de Amazon habilitada, que se incluye en las versiones 6.2.0 a 7.3.0, y quieres actualizar a una versión con OSS Store File Tracking, que está disponible en las versiones de EMR posteriores a la 7.3.0, sigue estos pasos:

1. En el clúster existente:
    1. Deshabilite la `hbase:storefile` tabla.
    2. Elimine la tabla `hbase:storefile`.
    3. Al rash`base:meta`.
    4. Espere a que se actualicen los metadatos.
2. En el nuevo clúster:
    1. Establezca el mismo directorio de Amazon S3 que el directorio raíz.
    2. Inicie el clúster con la `DefaultStoreFileTracker` implementación:

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

    3. A nivel de familia de tablas o columnas, utilice los siguientes comandos para cambiar el rastreador de archivos del almacén:
        a. Cambia el Store File Tracker de la tabla o familia de columnas de la tabla:

```
hbase> change_sft 't1','FILE'
hbase> change_sft 't2','cf1','FILE'
```

b. Cambia todos los registros de archivos almacenados de la tabla que coincidan con la expresión regular (regex) dada:

```
hbase> change_sft_all 't.*','FILE'
hbase> change_sft_all 'ns:.*','FILE'
hbase> change_sft_all 'ns:t.*','FILE'
```

HBase historial de versiones

En la siguiente tabla se muestra la versión HBase incluida en cada versión de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

HBase información sobre la versión

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.9.0	2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.8.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.7.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.6.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.5.0	2.5.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.4.0	2.5.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.3.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.2.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.1.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-7.0.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.15.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.14.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.13.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.12.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.11.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.11.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.10.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.10.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.9.1	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.9.0	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.8.1	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.8.0	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.7.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.36.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.6.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.35.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.5.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.4.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.3.1	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.3.0	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.2.1	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.2.0	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.1.1	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.1.0	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.0.1	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-6.0.0	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.34.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.33.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.33.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.32.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.32.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.31.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.31.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.30.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.30.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.30.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.29.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.28.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.28.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.27.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.27.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.26.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.25.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.24.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.24.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.23.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.23.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.22.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.2	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.21.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.20.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.20.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.17.2	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.17.1	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.17.0	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.16.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.16.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.15.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.15.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.2	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.14.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.13.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.13.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.12.3	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.2	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.12.1	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.0	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.11.4	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.11.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.11.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.10.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.10.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.9.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.9.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.8.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.8.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.7.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.7.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.6.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.6.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.4	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.5.3	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.2	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.5.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.4.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.4.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.3.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.3.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.3.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.3	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.2.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.2.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.1.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.1.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.0.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-5.0.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.6	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.9.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.9.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.9.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.8.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.8.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.8.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.4	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.7.3	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.2	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-regionserver, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.7.1	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.0	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	HBase Versión	Componentes instalados con HBase
emr-4.6.1	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.6.0	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Apache HCatalog

HCatalog es una herramienta que te permite acceder a las tablas del metaalmacén de Hive desde Pig, Spark SQL o aplicaciones personalizadas. MapReduce HCatalog tiene una interfaz REST y un cliente de línea de comandos que te permiten crear tablas o realizar otras operaciones. A continuación, escriba sus aplicaciones para acceder a las tablas mediante HCatalog bibliotecas. Para obtener más información, consulte Uso HCatalog. HCatalog se incluye en la versión de lanzamiento de Amazon EMR 4.4.0 y posteriores.

HCatalog en Amazon EMR, la versión 5.8.0 y posteriores admite el uso del Catálogo de datos de AWS Glue como metaalmacén de Hive. Para más información, consulte Uso del Catálogo de datos de AWS Glue como metaalmacén de Hive.

En la tabla siguiente, se muestra la versión HCatalog incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes con los que Amazon EMR se instala. HCatalog

Para ver la versión de los componentes instalados HCatalog en esta versión, consulte Versiones del componente de la versión 6.15.0.

HCatalog información de la versión de emr-6.15.0

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.15.0	HCatalog 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

En la tabla siguiente, se muestra la versión HCatalog incluida en la última versión de la serie 5.x de Amazon EMR, junto con los componentes con los que Amazon EMR se instala. HCatalog

Para ver la versión de los componentes instalados HCatalog en esta versión, consulte Versiones del componente de la versión 5.36.2.

HCatalog información de la versión de emr-5.36.2

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.36.2	HCatalog 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Temas

- Crear un clúster con HCatalog
- Usando HCatalog
- Ejemplo: crear una HCatalog tabla y escribir en ella utilizando Pig
- HCatalog historial de lanzamientos

Crear un clúster con HCatalog

Aunque HCatalog está incluido en el proyecto Hive, debes instalarlo como una aplicación propia.

Para lanzar un clúster HCatalog con la consola

El siguiente procedimiento crea un clúster con HCatalog Installed. Para obtener más información sobre la creación de clústeres con la consola, incluidas Opciones avanzadas, consulte Planificar y configurar clústeres en la Guía de administración de Amazon EMR.

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Create cluster (Crear clúster) para utilizar Quick Create (Creación rápida).
3. Para el campo Software Configuration (Configuración de software), elija Amazon Release Version emr-4.4.0 (Versión de Amazon emr-4.4.0) o una versión posterior.
4. En el campo Seleccionar aplicaciones, elija Todas las aplicaciones o. HCatalog
5. Seleccione las demás opciones que necesite y, a continuación, elija Create cluster (Crear clúster).

Para lanzar un clúster HCatalog mediante el AWS CLI

- Cree el clúster con el siguiente comando:

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Cluster with Hcat" --release-label emr-7.9.0 \
--applications Name=HCatalog --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Usando HCatalog

Se puede utilizar HCatalog en varias aplicaciones que utilizan la metatienda Hive. Los ejemplos que aparecen en esta sección muestran cómo crear una tabla y utilizarla en el contexto de Pig y Spark SQL.

Deshabilitación de la escritura directa al utilizar HCatalog HStorer

Cuando una aplicación utilice HCatStorer para escribir en una HCatalog tabla almacenada en Amazon S3, deshabilite la característica de escritura directa de Amazon EMR. Por ejemplo, deshabilite la escritura directa si utiliza el `STORE` comando Pig o está ejecutando trabajos de Sqoop que escriben en HCatalog tablas en Amazon S3. Puede deshabilitar la característica de escritura directa estableciendo la configuración de `mapred.output.direct.NativeS3FileSystem` y `mapred.output.direct.EmrFileSystem` en `false`. El siguiente ejemplo muestra cómo definir estas configuraciones utilizando Java.

```
Configuration conf = new Configuration();
conf.set("mapred.output.direct.NativeS3FileSystem", "false");
conf.set("mapred.output.direct.EmrFileSystem", "false");
```

Creación de una tabla mediante la HCat CLI y uso de dichos datos en Pig

Cree el siguiente script, `impressions.q`, en su clúster:

```
CREATE EXTERNAL TABLE impressions (
    requestBeginTime string, adId string, impressionId string, referrer string,
    userAgent string, userCookie string, ip string
  )
  PARTITIONED BY (dt string)
  ROW FORMAT
    serde 'org.apache.hive.hcatalog.data.JsonSerDe'
    with serdeproperties ( 'paths'='requestBeginTime, adId, impressionId, referrer,
 userAgent, userCookie, ip' )
  LOCATION 's3://[your region].elasticmapreduce/samples/hive-ads/tables/impressions/';
ALTER TABLE impressions ADD PARTITION (dt='2009-04-13-08-05');
```

Ejecute el script mediante la HCat CLI:

```
% hcat -f impressions.q
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-
log4j.properties
OK
Time taken: 4.001 seconds
OK
Time taken: 0.519 seconds
```

Abra el shell de Grunt y acceda a los datos en `impressions`:

```
% pig -useHCatalog -e "A = LOAD 'impressions' USING
 org.apache.hive.hcatalog.pig.HCatLoader();
B = LIMIT A 5;
dump B;"
<snip>
(1239610346000,m9nwdo67Nx6q2kI25qt5On7peICfUM,omkxkaRpNhGPDucAiBErSh1cs0MThC,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 6.0; FunWebProducts; GTB6; SLCC1; .NET CLR
 2.0.50727; Media Center PC
 5.0; .NET,wcVWWTascoPbGt6bdqDbuWTPPHgOPs,69.191.224.234,2009-04-13-08-05)
(1239611000000,NjriQjdODgWBKnkGJUP6GNTbDeK4An,AWtXPkfaWGOaNeL9OOsFU8Hcj6eLHt,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 5.1; GTB6; .NET CLR
 1.1.4322),OaMU1F2gE4CtADVHAbKjjRRks5kIgg,57.34.133.110,2009-04-13-08-05)
(1239610462000,Irpv3oiu0I5QNQiwSSTIshrLdo9cM1,i1LDq44LRSJF0hbmhB8Gk7k9gMWtBq,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239611007000,q2Awfnpe0JAvhInaIp0VGx9KTs0oPO,s3HvTflPB8JIE0IuM6hOEebWWpOtJV,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239610398000,c362vpAB0soPKGHRS43cj6TRwNeOGn,jeas5nXbQInGAgFB8jlkhnprN6cMw7,cartoonnetwork.com
 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; GTB6; .NET CLR
 1.1.4322),k96n5PnUmwHKfiUI0TFP0TNMfADgh9,51.131.29.87,2009-04-13-08-05)
7120 [main] INFO  org.apache.pig.Main  - Pig script completed in 7 seconds and 199
 milliseconds (7199 ms)
16/03/08 23:17:10 INFO pig.Main: Pig script completed in 7 seconds and 199 milliseconds
 (7199 ms)
```

Acceso a la tabla mediante Spark SQL

En este ejemplo se crea un Spark DataFrame a partir de la tabla creada en el primer ejemplo y se muestran las primeras 20 líneas:

```
% spark-shell --jars /usr/lib/hive-hcatalog/share/hcatalog/hive-hcatalog-core-1.0.0-
amzn-3.jar
<snip>
scala> val hiveContext = new org.apache.spark.sql.hive.HiveContext(sc);
scala> val df = hiveContext.sql("SELECT * FROM impressions")
scala> df.show()
<snip>
16/03/09 17:18:46 INFO DAGScheduler: ResultStage 0 (show at <console>:32) finished in
 10.702 s
16/03/09 17:18:46 INFO DAGScheduler: Job 0 finished: show at <console>:32, took
 10.839905 s
```

```
+---------------+--------------------+--------------------+-----------------
+--------------------+--------------------+-------------+---------------+
|requestbegintime|                adid|        impressionid|          referrer|
   useragent|          usercookie|            ip|              dt|
+---------------+--------------------+--------------------+-----------------
+--------------------+--------------------+-------------+---------------+
|   1239610346000|m9nwdo67Nx6q2kI25...|omkxkaRpNhGPDucAi...|cartoonnetwork.com|
Mozilla/4.0 (comp...|wcVWWTascoPbGt6bd...|69.191.224.234|2009-04-13-08-05|
|   1239611000000|NjriQjdODgWBKnkGJ...|AWtXPkfaWGOaNeL9O...|cartoonnetwork.com|
Mozilla/4.0 (comp...|OaMU1F2gE4CtADVHA...| 57.34.133.110|2009-04-13-08-05|
|   1239610462000|Irpv3oiu0I5QNQiwS...|i1LDq44LRSJF0hbmh...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239611007000|q2Awfnpe0JAvhInaI...|s3HvTflPB8JIE0IuM...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239610398000|c362vpAB0soPKGHRS...|jeas5nXbQInGAgFB8...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610600000|cjBTpruoaiEtqLuMX...|XwlohBSs8Ipxs1bRa...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610804000|Ms3eJHNAEItpxvimd...|4SIj4pGmgVLl625BD...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610872000|h5bccHX6wJReDi1jL...|EFAWIiBdVfnxwAMWP...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610365000|874NBpGmxNFfxEPKM...|xSvE4XtGbdtXPF2Lb...|cartoonnetwork.com|
Mozilla/5.0 (Maci...|eWDEVVUphlnRa273j...| 22.91.173.232|2009-04-13-08-05|
|   1239610348000|X8gISpUTSqh1A5reS...|TrFblGT99AgE75vuj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610743000|kbKreLWB6QVueFrDm...|kVnxx9Ie2i3OLTxFj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610812000|9lxOSRpEi3bmEeTCu...|1B2sff99AEIwSuLVV...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610876000|lijjmCf2kuxfBTnjL...|AjvufgUtakUFcsIM9...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610941000|t8t8trgjNRPIlmxuD...|agu2u2TCdqWP08rAA...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610490000|OGRLPVNGxiGgrCmWL...|mJg2raBUpPrC8OlUm...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610556000|OnJID12x0RXKPUgrD...|P7Pm2mPdW6wO8KA3R...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610373000|WflsvKIgOqfIE5KwR...|TJHd1VBspNcua0XPn...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610768000|4MJR0XxiVCU1ueXKV...|1OhGWmbvKf8ajoU8a...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610832000|gWIrpDiN57i3sHatv...|RNL4C7xPi3tdar2Uc...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
```

```
|   1239610789000|pTne9k62kJ14QViXI...|RVxJVIQousjxUVI3r...|        pixnet.net|
Mozilla/5.0 (Maci...|1bGOKiBD2xmui9OkF...| 33.176.101.80|2009-04-13-08-05|
+----------------+--------------------+--------------------+-----------------
+--------------------+--------------------+-------------+---------------+
only showing top 20 rows


scala>
```

Ejemplo: crear una HCatalog tabla y escribir en ella utilizando Pig

Puede crear una HCatalog tabla y utilizar Apache Pig para escribir en ella mediante HCat Storer utilizando una fuente de datos en Amazon S3. HCatalog requiere que deshabilite la escritura directa o la operación fallará silenciosamente. Defina las configuraciones `mapred.output.direct.NativeS3FileSystem` y `mapred.output.direct.EmrFileSystem` en `false` utilizando la clasificación `mapred-site` o bien manualmente desde el shell de Grunt. En el siguiente ejemplo, se muestra una tabla creada con la HCat CLI, seguida de comandos ejecutados en el intérprete de comandos de Grunt para rellenar la tabla a partir de un archivo de datos de ejemplo en Amazon S3.

Para ejecutar este ejemplo, conéctese al nodo maestro mediante SSH.

Cree un archivo de HCatalog script`wikicount.q`, con el siguiente contenido, para crear una HCatalog tabla denominada. `wikicount`

```
CREATE EXTERNAL TABLE IF NOT EXISTS wikicount(
col1 string,
col2 bigint
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\001'
STORED AS ORC
LOCATION 's3://amzn-s3-demo-bucket/hcat/wikicount';
```

Utilice un comando HCat CLI para ejecutar el script desde el archivo.

```
hcat -f wikicount.q
```

A continuación, inicie el shell de Grunt con la opción `-useHCatalog`, defina las configuraciones para deshabilitar la escritura directa, cargar datos desde una ubicación de S3 y, a continuación, escriba los resultados en la tabla wikicount.

```
pig -useHCatalog
SET mapred.output.direct.NativeS3FileSystem false;
SET mapred.output.direct.EmrFileSystem false;
A = LOAD 's3://support.elasticmapreduce/training/datasets/wikistats_tiny/' USING
 PigStorage(' ') AS (Site:chararray, page:chararray, views:int, total_bytes:long);
B = GROUP A BY Site;
C = FOREACH B GENERATE group as col1, COUNT(A) as col2;
STORE C INTO 'wikicount' USING org.apache.hive.hcatalog.pig.HCatStorer();
```

HCatalog historial de lanzamientos

En la tabla siguiente, se muestra la versión HCatalog incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

HCatalog información sobre la versión

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-s

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
		erver, hive-client, mariadb-s erver
emr-7.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-7.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.36.2	2.3.9.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.36.1	2.3.9.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.36.0	2.3.9.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.35.0	2.3.9.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Etiqueta de versión de Amazon EMR	HCatalog Versión	Componentes instalados con HCatalog
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Apache Hive

Hive es un almacén de datos y paquete de análisis de código abierto que se ejecuta sobre un clúster de Hadoop. Los scripts de Hive usan un lenguaje como SQL llamado Hive QL (lenguaje de consultas) que resume los modelos de programación y admite las interacciones típicas de un almacén de datos. Hive le permite evitar las complejidades de escribir trabajos de Tez basados en gráficos acíclicos dirigidos (DAGs) o MapReduce programas en un lenguaje informático de nivel inferior, como Java.

Hive amplía el paradigma de SQL incluyendo formatos de serialización. También puede personalizar el procesamiento de consultas creando un esquema de tabla acorde con sus datos, sin tocar los datos. Aunque SQL solo es compatible con tipos de valor primitivos (como fechas, números y cadenas), los valores de las tablas de Hive son elementos estructurados, por ejemplo, objetos JSON, cualquier tipo de datos definido por el usuario o cualquier función escrita en Java.

Para obtener más información acerca de Hive, consulte http://hive.apache.org.

En la siguiente tabla, se muestra la versión de Hive incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Hive.

Para ver la versión de los componentes instalados con Hive en esta versión, consulte Versiones de componentes de la versión 7.9.0.

Información sobre la versión de Hive para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-7.9.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
		rn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

En la tabla siguiente, se muestra la versión de Hive incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Hive.

Para ver la versión de los componentes instalados con Hive en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Hive para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.15.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

En la tabla siguiente, se muestra la versión de Hive incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Hive.

Para ver la versión de los componentes instalados con Hive en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Hive para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.36.2	Hive 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

A partir de Amazon EMR 5.18.0, puede utilizar el repositorio de artefactos de Amazon EMR para compilar el código de trabajos con las versiones exactas de las bibliotecas y dependencias que están disponibles con determinadas versiones de Amazon EMR. Para obtener más información, consulte Comprobación de dependencias mediante el repositorio de artefactos de Amazon EMR.

Temas

- Diferencias y consideraciones sobre Hive en Amazon EMR
- Configuración de un metaalmacén externo para Hive
- Uso del controlador JDBC de Hive
- Mejora del rendimiento de Hive
- Uso de Hive Live Long and Process (LLAP)

- Cifrado en Hive
- Historial de versiones de Hive

Diferencias y consideraciones sobre Hive en Amazon EMR

Diferencias entre Apache Hive en Amazon EMR y Apache Hive

En esta sección, se describen las diferencias entre Hive en Amazon EMR y las versiones predeterminadas de Hive disponibles en http://svn.apache. org/viewvc/hive/branches/.

Autorización de Hive

Amazon EMR admite Autorización de Hive para HDFS, pero no para EMRFS y Amazon S3. Los clústeres de Amazon EMR se ejecutan con la autorización desactivada de forma predeterminada.

Comportamiento de combinación de archivos de Hive con Amazon S3

Apache Hive combina archivos pequeños al final de un trabajo de solo mapeo si `hive.merge.mapfiles` es true y la combinación se activa solo si el tamaño de salida promedio del trabajo es inferior al ajuste `hive.merge.smallfiles.avgsize`. Hive en Amazon EMR tiene exactamente el mismo comportamiento si la ruta de salida final se encuentra en HDFS. Si la ruta de salida se encuentra en Amazon S3, se omite el parámetro `hive.merge.smallfiles.avgsize`. En esta situación, la tarea de combinación siempre se activa si `hive.merge.mapfiles` está definido en `true`.

Transacciones de ACID y Amazon S3

La versión 6.1.0 y posteriores de Amazon EMR admiten transacciones de ACID (atomicidad, coherencia, aislamiento y durabilidad) de Hive, por lo que cumplen con las propiedades de ACID de una base de datos. Con esta característica, puede ejecutar operaciones INSERT, UPDATE, DELETE y MERGE en tablas administradas por Hive con datos de Amazon Simple Storage Service (Amazon S3).

Hive Live Long and Process (LLAP)

La funcionalidad LLAP agregada en la versión 2.0 de Apache Hive predeterminado no se admite en la versión 2.1.0 de Hive en Amazon EMR versión 5.0.

La versión 6.0.0 y posteriores de Amazon EMR son compatibles con la funcionalidad Live Long and Process (LLAP) para Hive. Para obtener más información, consulte Uso de Hive LLAP.

Diferencias de Hive entre la versión de lanzamiento 4.x y 5.x de Amazon EMR

En esta sección, se abordan las diferencias que deben tenerse en cuenta antes de migrar una implementación de Hive desde la versión 1.0.0 de Hive en Amazon EMR versión 4.x a Hive 2.x en Amazon EMR versión 5.x.

Consideraciones y diferencias operativas

- Se ha agregado compatibilidad con Transacciones de ACID (atomicidad, uniformidad, aislamiento y durabilidad): se ha eliminado esta diferencia entre Hive 1.0.0 en Amazon EMR 4.x y Apache Hive predeterminado.
- Se han eliminado las escrituras directas en Amazon S3: se ha eliminado esta diferencia entre Hive 1.0.0 en Amazon EMR y el Apache Hive predeterminado. Hive 2.1.0 en Amazon EMR versión 5.x ahora crea, lee y escribe en archivos temporales almacenados en Amazon S3. En consecuencia, para leer y escribir en la misma tabla ya no tiene que crear una tabla temporal en el sistema de archivos HDFS local del clúster como solución. Si utiliza los buckets con control de versiones, asegúrese de administrar estos archivos temporales tal y como se describe a continuación.
- Administrar archivos temporales al utilizar buckets de Amazon S3 con control de versiones: cuando se ejecutan consultas de Hive donde el destino de los datos generados es Amazon S3, se crean muchos archivos y directorios temporales. Este es el nuevo comportamiento tal y como se describió anteriormente. Si utiliza buckets de S3 con control de versiones, estos archivos temporales desordenan Amazon S3 e incurren en costos si no se eliminan. Ajuste sus reglas de ciclo de vida de modo que los datos con prefijo `/_tmp` se eliminen tras un breve periodo de tiempo como, por ejemplo, cinco días. Consulte Especificación de una configuración del ciclo de vida para más información.
- Log4j actualizado a log4j 2: si utiliza log4j, es posible que tenga que cambiar su configuración de registro debido a esta actualización. Consulte Apache log4j 2 para obtener más información.

Diferencias y consideraciones de rendimiento

- Diferencias de rendimiento con Tez: con la versión 5.x de Amazon EMR, Tez es el motor de ejecución predeterminado de Hive en lugar de. MapReduce Tez proporciona un rendimiento mejorado para la mayoría de flujos de trabajo.
- Tablas con muchas particiones: las consultas que generan un gran número de particiones dinámicas podrían devolver errores y las consultas que seleccionan desde tablas con muchas

particiones podrían tardar más de lo esperado en ejecutarse. Por ejemplo, una selección desde 100 000 particiones podría tardar 10 minutos o más.

Características adicionales de Hive en Amazon EMR

Amazon EMR amplía Hive con nuevas características que admiten la integración de Hive con otros AWS servicios de, como la posibilidad de leer y escribir en Amazon Simple Storage Service (Amazon S3) y DynamoDB.

Variables en Hive

Puede incluir variables en sus scripts utilizando el signo de dólar y llaves.

```
add jar ${LIB}/jsonserde.jar
```

Los valores de estas variables se transfieren a Hive en la línea de comando con el parámetro `-d`, como en el siguiente ejemplo:

```
-d LIB=s3://elasticmapreduce/samples/hive-ads/lib
```

También puede transferir los valores en pasos que ejecutan scripts de Hive.

Para transferir valores de variable en pasos de Hive utilizando la consola

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Create cluster.
3. En la sección Steps (Pasos), en Add Step (Añadir paso), elija Hive Program (Programa de Hive) en la lista y Configure and add (Configurar y añadir).
4. En el cuadro de diálogo Add Step (Añadir paso), especifique los parámetros utilizando la tabla siguiente como guía y, a continuación, elija Add (Añadir).

Campo	Acción
Script S3 location*	Especifique el URI donde reside el script en Amazon S3. El valor debe tener el formato*BucketName* //. *path ScriptName* Por ejemplo: `s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q` .
Input S3 location	De manera opcional, especifique el URI en el que se encuentra n los archivos de entrada en Amazon S3. El valor debe tener el formato*BucketName* /*path*/. Si se especifica, se transferirá al script de Hive como un parámetro denominado `INPUT`. Por ejemplo: `s3://elasticmapreduce/samples/hive-ads/tables/` .
Output S3 location	De manera opcional, especifique el URI donde desee la salida almacenada en Amazon S3. El valor debe tener el formato*BucketName* /*path*. Si se especifica, se transferirá al script de Hive como un parámetro denominado OUTPUT. Por ejemplo: `s3://amzn-s3-demo-bucket/hive-ads/output/` .
Argumentos	Opcionalmente, introduzca una lista de argumentos (cadenas separadas por espacios) que transferir a Hive. Si ha definido una variable de ruta en el script de Hive denominada ${SAMPLE}, por ejemplo: `CREATE EXTERNAL TABLE logs (requestBeginTime STRING, requestEndTime STRING, hostname STRING) PARTITIONED BY (dt STRING) \ ROW FORMAT serde 'com.amazon.elasticmapreduce.JsonSerde' WITH SERDEPROPERTIES ( 'paths'='requestBeginTime, requestEndTime, hostname' ) LOCATION '${SAMPLE}/tables/impressions';` Para pasar un valor a la variable, escriba lo siguiente en la ventana Arguments (Argumentos): `-d SAMPLE=s3://elasticmapreduce/samples/hive-ads/` .

Campo	Acción
Action on Failure	Esto determina qué hace el clúster en respuesta a los errores. Los valores posibles para esta configuración son: • Terminate cluster (Terminar clúster): si el paso genera un error, terminar el clúster. Si el clúster tiene la protección de terminación habilitada Y keep-alive habilitado, no se terminará. • Cancel and wait (Cancelar y esperar): si el paso genera un error, cancelar los pasos restantes. Si el clúster tiene keep-alive habilitado, el clúster no termina. • Continue (Continuar): si el paso genera un error, continuar en el paso siguiente.

5. Seleccione los valores que sean necesarios y elija Create cluster (Crear clúster).

Para pasar los valores de las variables a los pasos de Hive, utilice el AWS CLI

Para pasar valores variables a los pasos de Hive mediante el AWS CLI, utilice el `--steps` parámetro e incluya una lista de argumentos.

•

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hive Name=Pig --use-default-roles --ec2-attributes
 KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,INPUT=s3://
elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/hive-
ads/output/,-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Para transferir valores de variable en pasos de Hive utilizando el SDK de Java

- El siguiente ejemplo muestra cómo transferir variables a pasos con el SDK. Para obtener más información, consulte Class StepFactory in the AWS SDK para Java API Reference.

```
StepFactory stepFactory = new StepFactory();

   StepConfig runHive = new StepConfig()
     .withName("Run Hive Script")
     .withActionOnFailure("TERMINATE_JOB_FLOW")
     .withHadoopJarStep(stepFactory.newRunHiveScriptStep(“s3://amzn-s3-demo-bucket/
script.q”,
      Lists.newArrayList(“-d”,”LIB= s3://elasticmapreduce/samples/hive-ads/lib”));
```

Consultas de Hive de Amazon EMR para asignar esquemas parciales de DynamoDB

Hive de Amazon EMR ofrece la máxima flexibilidad a la hora de consultar las tablas de DynamoDB permitiéndole especificar un subconjunto de columnas en el que puede filtrar los datos, en lugar de exigir que la consulta incluya todas las columnas. Esta técnica de consulta de esquema parcial resulta eficaz cuando tenga un esquema de base de datos disperso y desee filtrar registros en función de algunas columnas, como, por ejemplo, filtrado de marcas temporales.

El siguiente ejemplo muestra cómo utilizar una consulta de Hive para:

- Crear una tabla de DynamoDB.
- Seleccione un subconjunto de elementos (filas) en DynamoDB y reduzca aún más los datos en ciertas columnas.
- Copie los datos resultantes a Amazon S3.

```
DROP TABLE dynamodb;
DROP TABLE s3;

CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, fullColumn
 map<String, String>)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

```
CREATE EXTERNAL TABLE s3(map<String, String>)
     ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
     LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3 SELECT item fullColumn FROM dynamodb WHERE recordTimeStamp <
 "2012-01-01";
```

En la siguiente tabla se muestra la sintaxis de consultas para seleccionar cualquier combinación de elementos desde DynamoDB.

Ejemplo de consulta	Descripción del resultado
SELECCIONA* DE*table_name* ;	Selecciona todos los elementos (filas) de una tabla determinada e incluye datos de todas las columnas disponibles para dichos elementos.
SELECCIONE * DESDE *table_name* DONDE *field_name* =*value*;	Selecciona algunos elementos (filas) de una tabla determinada e incluye datos de todas las columnas disponibles para dichos elementos.
SELECCIONA*column1_name* ,*column2_name* , *column3_name* DE*table_name* ;	Selecciona todos los elementos (filas) de una tabla determinada e incluye datos de algunas columnas disponibles para dichos elementos.
SELECCIONE*column1_name* ,*column2_name* , *column3_name* DESDE *table_name* DONDE *field_name* =*value*;	Selecciona algunos elementos (filas) de una tabla determinada e incluye datos de algunas columnas disponibles para dichos elementos.

Copia de datos entre tablas de DynamoDB de distintas regiones de AWS

Hive de Amazon EMR proporciona una propiedad `dynamodb.region` que puede definir en cada tabla de DynamoDB. Cuando `dynamodb.region` se define de forma distinta en dos tablas, todos los datos que copie entre las tablas se dan automáticamente entre las regiones especificadas.

El siguiente ejemplo muestra cómo crear una tabla de DynamoDB con un script de Hive que define la propiedad `dynamodb.region`:

> **Note**
>
> Las propiedades de región por tabla anulan las propiedades de Hive globales.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.region" = "eu-west-1",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

Definir los valores de rendimiento de DynamoDB por tabla

Hive de Amazon EMR le permite definir DynamoDB readThroughputPercent y la writeThroughputPercent configuración de cada tabla en la definición de la tabla. El siguiente script de Hive de Amazon EMR muestra cómo definir los valores de rendimiento. Para más información sobre los valores de rendimiento de DynamoDB, consulte Especificación de los requisitos de lectura y escritura para las tablas.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".4",
     "dynamodb.throughput.write.percent" = "1.0",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

Configuración de un metaalmacén externo para Hive

Hive registra la información del metaalmacén de manera predeterminada en una base de datos de MySQL que se encuentra en el sistema de archivos del nodo principal. El metaalmacén contiene una descripción de la tabla y los datos subyacentes a partir de los que se creó, incluidos los nombres de particiones, tipos de datos, etc. Cuando un clúster termina, todos los nodos del clúster se cierran, incluido el nodo principal. Cuando esto ocurre, se pierden los datos locales porque los sistemas

de archivos del nodo usan almacenamiento efímero. Si necesita que se conserven los datos del metaalmacén, debe crear un metaalmacén externo que exista fuera del clúster.

Dispone de dos opciones para crear un metaalmacén externo:

- AWS Catálogo de datos de Glue (solo Amazon EMR versión 5.8.0 o posteriores).

  Para obtener más información, consulte Se ha agregado compatibilidad para utilizar el Catálogo de datos de AWS Glue como metaalmacén predeterminado de Hive.
- Amazon RDS o Amazon Aurora.

  Para obtener más información, consulte Uso de una base de datos MySQL externa o Amazon Aurora.

> **Note**
>
> Si utiliza Hive 3 y encuentra demasiadas conexiones con el metaalmacén de Hive, configure el parámetro `datanucleus.connectionPool.maxPoolSize` para que tenga un valor menor o aumente el número de conexiones que puede gestionar el servidor de base de datos. El aumento en el número de conexiones se debe a la forma en que Hive calcula el número máximo de conexiones JDBC. Para calcular el valor óptimo de rendimiento, consulte Hive Configuration Properties.

Se ha agregado compatibilidad para utilizar el Catálogo de datos de AWS Glue como metaalmacén predeterminado de Hive

Con la versión 5.8.0 o posterior de Amazon EMR, puede configurar Hive para utilizar el Catálogo de datos de AWS Glue como metaalmacén. Recomendamos esta configuración cuando se necesita un metaalmacén persistente o un metaalmacén compartido por diferentes clústeres, servicios, aplicaciones o cuentas de AWS .

AWS Glue es un servicio completamente administrado de ETL (extracción, transformación y carga) con el que resulta más rentable y sencillo categorizar los datos, limpiarlos, enriquecerlos y moverlos de manera fiable entre distintos almacenes de datos. El Catálogo de datos de AWS Glue proporciona un repositorio de metadatos que unifica toda una variedad de orígenes y formatos de datos, e integra no solo Amazon EMR, sino también Amazon RDS, Amazon Redshift, Redshift, Redshift Spectrum, Athena y todas las aplicaciones compatibles con el metaalmacén de Apache Hive. AWS

Los rastreadores de Glue pueden deducir automáticamente el esquema de los datos de origen en Amazon S3 y almacenar los metadatos asociados en el Catálogo de datos. Para más información acerca del Catálogo de datos, consulte Cómo rellenar el Catálogo de datos de AWS Glue en la Guía para desarrolladores de AWS Glue.

Se aplican cargos separados para AWS Glue. Se paga una tarifa mensual para almacenar los metadatos en el Catálogo de datos y acceder a ellos; una tarifa por hora facturada por minuto para los trabajos de ETL de AWS Glue y el tiempo de ejecución del rastreador; y una tarifa por hora facturada por minuto por cada punto de conexión de desarrollo aprovisionado. El Catálogo de datos le permite almacenar hasta un millón de objetos sin cargo alguno. Si almacena más de un millón de objetos, se le cobrará 1 USD por cada 100 000 objetos que sobrepasen esa cifra. En el Catálogo de datos, se consideran objetos las tablas, las particiones y las bases de datos. Para obtener más información, consulte Precios de Glue.

> **Important**
>
> Si ha creado tablas con Amazon Athena o Amazon Redshift Spectrum antes del 14 de agosto de 2017, las bases de datos y las tablas se almacenan en un catálogo administrado por Athena, que es independiente del Catálogo de datos de Glue. AWS Para integrar Amazon EMR con estas tablas, debe actualizarse al Catálogo de datos de AWS Glue. Para obtener más información, consulte Actualización al catálogo de datos de AWS Glue en la Guía del usuario de Amazon Athena.

Especificación del Catálogo de datos de AWS Glue como metaalmacén

Puede especificar el Catálogo de datos de AWS Glue como metaalmacén mediante la AWS Management Console AWS CLI, la y la API de Amazon EMR. Cuando utiliza la CLI o la API, utiliza la clasificación de configuración de Hive para especificar el catálogo de datos. Además, con Amazon EMR 5.16.0 y versiones posteriores, puede utilizar la clasificación de configuración para especificar un Catálogo de datos en otra. Cuenta de AWS Cuando utiliza la consola, puede especificar el Catálogo de datos mediante Opciones avanzadas u Opciones rápidas.

Console

Para especificar el Catálogo de datos de AWS Glue como metaalmacén de Hive con la consola

1. Inicie sesión en la y abra la AWS Management Console consola de Amazon EMR en /emr en https://console.aws.amazon.com /emr.

2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Paquete de aplicaciones, elija Core Hadoop o HBaseCustom. Si personaliza el clúster, asegúrese de seleccionar Hive o HCatalog como una de sus aplicaciones.
4. En Configuración del Catálogo de datos de AWS Glue, seleccione la casilla Usar para metadatos de la tabla de Hive.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

CLI

Para especificar el Catálogo de datos de AWS Glue como metaalmacén de Hive con la AWS CLI

Para obtener más información sobre cómo especificar una clasificación de configuración con la AWS CLI y la API de EMR, consulte. Configuración de aplicaciones

- Especifique el valor de `hive.metastore.client.factory.class` utilizando la clasificación de configuración `hive-site` tal y como se muestra en el siguiente ejemplo:

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

En las versiones 5.28.0, 5.28.1, 5.29.0 o 6.x de EMR, si va a crear un clúster con el Catálogo de datos de AWS Glue como metaalmacén, establezca el valor de. `hive.metastore.schema.verification` `false` Esto impide que Hive HCatalog valide el esquema de metastore con MySQL. Sin esta configuración, el grupo de instancias principal se suspenderá después de la reconfiguración en Hive o. HCatalog

```
[
  {
```

```
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false"
    }
  }
]
```

Si ya tiene un clúster en la versión de EMR 5.28.0, 5.28.1 o 5.29.0 de EMR, puede establecer el grupo de instancias principal `hive.metastore.schema.verification` como `false` con la siguiente información:

```
    Classification = hive-site
    Property       = hive.metastore.schema.verification
    Value          = false
```

Para especificar un Catálogo de datos en otra AWS cuenta de, agregue la `hive.metastore.glue.catalogid` propiedad tal y como se muestra en el siguiente ejemplo. Sustituya *`acct-id`* por la cuenta de AWS del Catálogo de datos.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Permisos de IAM

El perfil de EC2 instancia de un clúster debe tener permisos de IAM para las acciones de AWS Glue. Además, si se habilita el cifrado para objetos del Catálogo de datos de AWS Glue, el rol también debe poder cifrar, descifrar y generar la AWS KMS key utilizada para el cifrado.

Permisos para las acciones de AWS Glue

Si utiliza el perfil de EC2 instancia predeterminado para Amazon EMR, no se requiere ninguna acción. La política `AmazonElasticMapReduceforEC2Role` gestionada que se adjunta a la `EMR_EC2_DefaultRole` permite todas las acciones de AWS Glue necesarias. Sin embargo, si especifica permisos y un perfil de EC2 instancia personalizados, debe configurar las acciones de AWS Glue adecuadas. Utilice la política administrada `AmazonElasticMapReduceforEC2Role` como punto de partida. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia) en la Guía de administración de Amazon EMR.

Permisos para cifrar y descifrar el Catálogo de datos de AWS Glue

El perfil de instancia necesita permiso para cifrar y descifrar los datos con su clave. No es necesario configurar estos permisos si se cumplen las dos instrucciones siguientes:

- Para habilitar el cifrado de los objetos del Catálogo de datos de AWS Glue, utilice claves administradas para AWS Glue.
- Utilice un clúster que esté en la Cuenta de AWS misma que el Catálogo de datos de AWS Glue.

De lo contrario, debe agregar la siguiente instrucción a la política de permisos adjunta al perfil de EC2 instancia.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
```

```
            }
        ]
    }
]
```

Para más información sobre el cifrado del Catálogo de datos de AWS Glue, consulte Cifrado del Catálogo de datos en la Guía para desarrolladores de AWS Glue.

Permisos basados en recursos

Si utiliza AWS Glue junto con Hive, Spark o Presto en Amazon EMR, AWS Glue admite políticas basadas en recursos para controlar el acceso a los recursos del Catálogo de datos. Estos recursos incluyen bases de datos, tablas, conexiones y funciones definidas por el usuario. Para obtener más información, consulte Políticas de recursos de AWS Glue en la Guía para desarrolladores de AWS Glue.

Cuando utilice políticas basadas en recursos para limitar el acceso a AWS Glue desde Amazon EMR, la entidad principal que especifique en la política de permisos debe ser el ARN del rol asociado al perfil de EC2 instancia que se especifica al crear un clúster. Por ejemplo, para una política basada en recursos adjunta a un catálogo, puede especificar el ARN del rol *EMR_EC2_DefaultRole* como rol de servicio predeterminado para las EC2 instancias del clúster`Principal`, con el formato que se muestra en el siguiente ejemplo:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*Puede ser diferente del ID de la cuenta de AWS Glue. Esto permite el acceso desde clústeres de EMR en diferentes cuentas. Puede especificar varias entidades principales, cada una de ellas desde una cuenta diferente.

Consideraciones a la hora de utilizar el Catálogo de datos de AWS Glue

Tenga en cuenta los siguientes elementos al utilizar el Catálogo de datos de AWS Glue como metaalmacén con Hive:

- No se admite la adición de elementos auxiliares JARs mediante el shell de Hive. Como solución alternativa, utilice la clasificación de `hive-site` configuración para establecer la `hive.aux.jars.path` propiedad, que añade el elemento auxiliar JARs a la ruta de clases de Hive.
- No se admiten las transacciones de Hive.

- No se admite el cambio de nombre de las tablas desde AWS Glue.
- Al crear una tabla de Hive sin especificar un valor de `LOCATION`, los datos de la tabla se almacenan en la ubicación especificada por la propiedad `hive.metastore.warehouse.dir`. De manera predeterminada, se trata de una ubicación en HDFS. Si otro clúster necesita acceder a la tabla, se produce un error a menos que tenga los permisos adecuados para el clúster que creó la tabla. Además, dado que el almacenamiento en HDFS es transitorio, si el clúster termina, se pierden los datos de la tabla y, por lo tanto, será necesario volver a crearla. Se recomienda que especifique un valor de a `LOCATION` en Amazon S3 cuando cree una tabla de Hive mediante AWS Glue. Como alternativa, puede utilizar la clasificación de configuración `hive-site` para especificar una ubicación en Amazon S3 para `hive.metastore.warehouse.dir`, que se aplica a todas las tablas de Hive. Si se crea una tabla en una ubicación de HDFS y el clúster que la creó sigue ejecutándose, puede actualizar la ubicación de la tabla a Amazon S3 desde AWS Glue. Para obtener más información, consulte Cómo trabajar con tablas en la consola de AWS Glue en la Guía para desarrolladores de AWS Glue.
- No se admiten valores de partición que contengan comillas y apóstrofes, por ejemplo, `PARTITION (owner="Doe's").`
- Las estadísticas de columnas son compatibles con emr-5.31.0 y versiones posteriores.
- No se admite el uso de la autorización de Hive. Como alternativa, considere la posibilidad de utilizar políticas basadas en recursos de AWS Glue. Para más información, consulte Uso de políticas basadas en recursos para el acceso de Amazon EMR a Catálogo de datos de Glue AWS.
- No se admiten las restricciones de Hive.
- No se admite la optimización basada en los costos en Hive.
- No se admite el uso de la configuración `hive.metastore.partition.inherit.table.properties`.
- No se admite el uso de las siguientes constantes de metaalmacén: `BUCKET_COUNT, BUCKET_FIELD_NAME, DDL_TIME, FIELD_TO_DIMENSION, FILE_INPUT_FORMAT, FILE_OUTPUT_FORMAT, HIVE_FILTER_FIELD_LAST_ACCESS, HIVE_FILTER_FIELD_OWNER, HIVE_FILTER_FIELD_PARAMS, IS_ARCHIVED, META_TABLE_COLUMNS, META_TABLE_COLUMN_TYPES, META_TABLE_DB, META_TABLE_LOCATION, META_TABLE_NAME, META_TABLE_PARTITION_COLUMNS, META_TABLE_SERDE, META_TABLE_STORAGE, ORIGINAL_LOCATION`.
- Cuando utiliza una expresión de predicado, los valores explícitos deben estar en el lado derecho del operador de comparación o las consultas podrían devolver errores.
  - Correcto: `SELECT * FROM mytable WHERE time > 11`

  - Incorrecto: `SELECT * FROM mytable WHERE 11 > time`
- Las versiones 5.32.0 y 6.3.0 y posteriores de Amazon EMR admiten el uso de funciones definidas por el usuario () en las expresiones de predicados. UDFs Si utiliza versiones anteriores, es posible que sus consultas presenten errores debido a la forma en que Hive intenta optimizar la ejecución de las consultas.
- No se admiten las tablas temporales.
- Se recomienda crear tablas con aplicaciones a través de Amazon EMR en lugar de crearlas directamente con Glue AWS . Al crear una tabla con AWS Glue, es posible que falten campos obligatorios y se produzcan excepciones en las consultas.
- En la versión 5.20.0 o posteriores de EMR, la eliminación de particiones en paralelo está habilitada automáticamente para Spark y Hive cuando se utiliza el Catálogo de datos de AWS Glue como metaalmacén. Este cambio reduce significativamente el tiempo de planificación de consultas al ejecutar varias solicitudes en paralelo para recuperar particiones. El número total de segmentos que se pueden ejecutar simultáneamente oscila entre 1 y 10. El valor predeterminado es 5, que es la configuración recomendada. Para cambiarlo, especifique la propiedad `aws.glue.partition.num.segments` en la clasificación de la configuración `hive-site`. Si se produce una limitación, puede desactivar la característica cambiando el valor a 1. Para más información, consulte Estructura Segment de AWS Glue.

Uso de una base de datos MySQL externa o Amazon Aurora

Para utilizar una base de datos MySQL externa o Amazon Aurora como metaalmacén de Hive, debe invalidar los valores de configuración predeterminados del metaalmacén de Hive para especificar la ubicación de la base de datos externa, ya sea en una instancia MySQL de Amazon RDS o en un Postgre de Amazon Aurora. SQLinstance

> **Note**
>
> Hive no admite ni impide el acceso de escritura simultánea a las tablas del metaalmacén. Si comparte información del metaalmacén entre dos clústeres, debe asegurarse de no escribir en la misma tabla de metaalmacén simultáneamente, a menos que se escriba en distintas particiones de la misma tabla.

El siguiente procedimiento muestra cómo omitir los valores de configuración predeterminados para la ubicación del metaalmacén de Hive e iniciar un clúster con la ubicación de metaalmacén reconfigurada.

Para crear un metaalmacén situado fuera del clúster de EMR

1. Cree una base de datos de MySQL o Aurora PostgreSQL. Si utiliza PostgreSQL, debe configurarlo después de aprovisionar el clúster. Solo se admite MySQL en la creación del clúster. Para obtener información sobre las diferencias entre Aurora MySQL y Aurora PostgreSQL, consulte Información general de Amazon Aurora MySQL y Uso de Amazon Aurora PostgreSQL. Para obtener información acerca de cómo crear una base de datos de Amazon RDS, consulte https://aws.amazon.com/rds/.
2. Modifique sus grupos de seguridad para permitir las conexiones JDBC entre su base de datos y el grupo de seguridad ElasticMapReduce-Master. Para más información sobre cómo modificar los grupos de seguridad para el acceso, consulte Uso de grupos de seguridad administrados por Amazon EMR.
3. Defina los valores de configuración de JDBC en `hive-site.xml`:

   > **Important**
   >
   > Si proporciona información confidencial, como contraseñas, a la API de configuración de Amazon EMR, esta información se muestra para aquellas cuentas que dispongan de permisos suficientes. Si le preocupa que esta información se pueda mostrar a otros usuarios, cree el clúster con una cuenta de administrador y limite a otros usuarios (usuarios de IAM o usuarios con credenciales delegadas) el acceso a los servicios del clúster creando un rol que deniegue de forma explícita los permisos para la clave de API `elasticmapreduce:DescribeCluster`.

   a. Cree un archivo de configuración llamado `hiveConfiguration.json` que contenga las modificaciones realizadas en `hive-site.xml` tal y como se muestra en el siguiente ejemplo.

      Sustituya *`hostname`* la dirección de DNS de la instancia de Amazon RDS que ejecuta la base de datos *`username`* y *`password`* por las credenciales de la base de datos. Para más información sobre la conexión a instancias de base de datos de MySQL y Aurora, consulte Conexión a una instancia de base de datos que ejecuta el motor de base de datos

de MySQL y Conexión a un clúster de base de datos de Athena en la Guía del usuario de Amazon RDS. `javax.jdo.option.ConnectionURL` es la cadena de conexión de JDBC para un metaalmacén de JDBC. `javax.jdo.option.ConnectionDriverName` es el nombre de clase de controlador para un metaalmacén de JDBC.

Amazon EMR instala los controladores de JDBC de MySQL.

La propiedad value no puede contener espacios ni saltos de línea. Debe aparecer todo en una única línea.

```
[
    {
      "Classification": "hive-site",
      "Properties": {
        "javax.jdo.option.ConnectionURL": "jdbc:mysql://hostname:3306/hive?
createDatabaseIfNotExist=true",
        "javax.jdo.option.ConnectionDriverName": "org.mariadb.jdbc.Driver",
        "javax.jdo.option.ConnectionUserName": "username",
        "javax.jdo.option.ConnectionPassword": "password"
      }
    }
  ]
```

b. Haga referencia al `hiveConfiguration.json` archivo cuando cree el clúster, como se muestra en el siguiente AWS CLI comando. En este comando, el archivo se almacena localmente, aunque también puede cargar el archivo en Amazon S3 y hacer referencia a él allí; por ejemplo, `s3://DOC-EXAMPLE-BUCKET/hiveConfiguration.json`.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=Hive --configurations file://hiveConfiguration.json --use-
default-roles
```

4. Conéctese al nodo principal en su clúster.

   Para más información, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR.

5. Cree sus tablas de Hive especificando la ubicación en Amazon S3 y escribiendo un comando similar al siguiente:

```
CREATE EXTERNAL TABLE IF NOT EXISTS table_name
(
key int,
value int
)
LOCATION s3://DOC-EXAMPLE-BUCKET/hdfs/
```

6. Añada el script de Hive al clúster en ejecución.

El clúster de Hive se ejecuta utilizando el metaalmacén ubicado en Amazon RDS. Lance todos los clústeres de Hive adicionales que comparten este metaalmacén especificando la ubicación del metaalmacén.

Uso del controlador JDBC de Hive

Puede usar herramientas de inteligencia empresarial populares MicroStrategy, como Microsoft Excel y Tableau con Amazon EMRQlikView, para explorar y visualizar sus datos. Muchas de estas herramientas requieren un controlador de conectividad de bases de datos Java (JDBC) o un controlador de conectividad de bases de datos abiertas (ODBC). Amazon EMR admite tanto la conectividad JDBC como ODBC.

El ejemplo siguiente muestra cómo usar SQL Workbench/J como un cliente SQL para conectarse a un clúster de Hive con Amazon EMR. Para ver otros controladores, consulte Utilizar herramientas de inteligencia empresarial con Amazon EMR.

Antes de instalar y trabajar con SQL Workbench/J, descargue el paquete del controlador e instale el controlador. Los controladores que se incluyen en el paquete admiten las versiones de Hive disponibles en la versión 4.0 y posteriores de Amazon EMR. Para ver las notas de la versión y documentación detallada, consulte la documentación de PDF incluida en el paquete.

- Descarga del paquete de controladores JDBC de Hive más reciente

http://awssupportdatasvcs.com/bootstrap-actions/Simba/latest/

- Versiones anteriores del controlador JDBC de Hive

  http://awssupportdatasvcs.com/bootstrap-actions/Simba/

Para instalar y configurar SQL Workbench

1. Descargue el cliente de SQL Workbench/J para su sistema operativo en http://www.sql-workbench.net/downloads.html.
2. Instale SQL Workbench/J. Para obtener más información, consulte Installing and starting SQL Workbench/J en el manual del usuario de SQL Workbench/J.
3. Usuarios de Linux, Unix, Mac OS X: en una sesión de terminal, cree un túnel de SSH al nodo principal del clúster utilizando el siguiente comando. *master-public-dns-name*Sustitúyalo por el nombre DNS público del nodo maestro y *path-to-key-file* por la ubicación y el nombre de archivo de tu archivo de clave EC2 privada de Amazon (`.pem`).

   ```
   ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000
    hadoop@master-public-dns-name
   ```

   Usuarios de Windows: en una sesión de PuTTY, cree un túnel SSH al nodo maestro de su clúster (mediante el reenvío de puertos locales) con `10000` para Puerto de origen y *master-public-dns-name*`:10000` para Destino. Sustituya *master-public-dns-name* por el nombre DNS público del nodo principal.
4. Añadir el controlador JDBC a SQL Workbench.

   a. En el cuadro de diálogo Select Connection Profile (Seleccionar perfil de conexión), haga clic en Manage Drivers (Administrar controladores).

   b. Haga clic en el icono Create a new entry (Crear una entrada nueva) (página en blanco).

   c. En el campo Name (Nombre), escriba **`Hive JDBC`**.

   d. En Library (Biblioteca), haga clic en el icono Select the JAR file(s) (Seleccionar los archivos JAR).

   e. Navegue hasta la ubicación que contiene los controladores extraídos. Seleccione los controladores que se incluyen en la versión del paquete de controladores JDBC que ha descargado y haga clic en Abrir.

Por ejemplo, su paquete de controladores JDBC puede incluir lo siguiente. JARs

```
hive_metastore.jar
hive_service.jar
HiveJDBC41.jar
libfb303-0.9.0.jar
libthrift-0.9.0.jar
log4j-1.2.14.jar
ql.jar
slf4j-api-1.5.11.jar
slf4j-log4j12-1.5.11.jar
TCLIServiceClient.jar
zookeeper-3.4.6.jar
```

f. En el cuadro de diálogo Please select one driver (Seleccione un controlador), seleccione `com.amazon.hive.jdbc41.HS2Driver`, seguido de OK (Aceptar).

5. Cuando vuelva al cuadro de diálogo Manage Drivers (Administrar controladores), compruebe que el campo Classname (Nombre de clase) esté relleno y seleccione OK (Aceptar).
6. Cuando vuelva al cuadro de diálogo Select Connection Profile (Seleccionar perfil de conexión), compruebe que el campo Driver (Controlador) esté definido en Hive JDBC (JDBC de Hive) y proporcione la siguiente cadena de conexión de JDBC en el campo URL: `jdbc:hive2://localhost:10000/default`.
7. Seleccione OK (Aceptar) para conectarse. Una vez se haya completado la conexión, los detalles de conexión en la parte superior de la ventana de SQL Workbench/J.

Para obtener más información sobre el uso de Hive y la interfaz JDBC, consulte la documentación Hive in Apache HiveClientJDBCInterfaceHive.

Mejora del rendimiento de Hive

Amazon EMR ofrece características para ayudar a optimizar el rendimiento cuando se utiliza Hive para consultar, leer y escribir datos guardados en Amazon S3.

S3 Select puede mejorar el rendimiento de las consultas para archivos JSON y CSV en algunas aplicaciones “delegando” el procesamiento en Amazon S3.

El archivador optimizado para EMRFS S3 es una alternativa a la OutputCommitterclase, que elimina las operaciones de enumeración y cambio de nombre para mejorar el rendimiento al escribir archivos en Amazon S3 mediante EMRFS.

Temas

- Habilitación del confirmador optimizado Hive EMRFS S3
- Uso de S3 Select con Hive para mejorar el rendimiento
- Optimización de MSCK

Habilitación del confirmador optimizado Hive EMRFS S3

El confirmador optimizado de Hive EMRFS S3 es una forma alternativa mediante la cual EMR Hive escribe archivos para consultas de inserción cuando se utiliza EMRFS. El confirmador elimina las operaciones de enumeración y cambio de nombre hechas en Amazon S3 y mejora el rendimiento de la aplicación. La característica está disponible a partir de EMR 5.34 y EMR 6.5.

Habilitación del confirmador

Si desea habilitar EMR Hive para que se utilice `HiveEMRFSOptimizedCommitter` para confirmar datos de forma predeterminada para todas las tablas externas y administradas por Hive, utilice la siguiente configuración `hive-site` en los clústeres de EMR 6.5.0 o EMR 5.34.0.

```
[
   {
      "classification": "hive-site",
      "properties": {
         "hive.blobstore.use.output-committer": "true"
      }
   }
]
```

> **Note**
>
> No active esta característica si `hive.exec.parallel` se ha establecido en `true`.

Limitaciones

Se aplican las siguientes restricciones básicas a las etiquetas:

- No se admite la posibilidad de que Hive combine archivos pequeños automáticamente. La lógica de confirmación predeterminada de Hive se utilizará incluso cuando el confirmador optimizado esté activado.
- No se admiten las tablas ACID de Hive. La lógica de confirmación predeterminada de Hive se utilizará incluso cuando el confirmador optimizado esté activado.
- La nomenclatura de los nombres de los archivos escritos ha cambiado de `<task_id>_<attempt_id>_<copy_n>` de Hive a `<task_id>_<attempt_id>_<copy_n>_<query_id>`. Por ejemplo, un archivo llamado `s3://warehouse/table/partition=1/000000_0` se cambiará a `s3://warehouse/table/partition=1/000000_0-hadoop_20210714130459_ba7c23ec-5695-4947-9d98-8a40ef759222-1`. `query_id` aquí es una combinación del nombre de usuario, la marca de tiempo y el UUID.
- Cuando las particiones personalizadas se encuentran en diferentes sistemas de archivos (HDFS, S3), esta característica se deshabilita automáticamente. La lógica de confirmación predeterminada de Hive se utilizará cuando se active.

Uso de S3 Select con Hive para mejorar el rendimiento

> **Important**
>
> Amazon S3 Select ya no está disponible para los nuevos clientes. Los clientes actuales de Amazon S3 Select pueden seguir utilizando la característica de la forma habitual. Más información

Con la versión 5.18.0 y posteriores de Amazon EMR, puede usar S3 Select con Hive en Amazon EMR. S3 Select permite que las aplicaciones recuperen únicamente un subconjunto de datos de un objeto. Para Amazon EMR, el clúster “delega” en Amazon S3 el trabajo de computación de filtrar grandes conjuntos de datos para su procesamiento, lo que puede mejorar el rendimiento en algunas aplicaciones y reduce la cantidad de datos transferidos entre Amazon EMR y Amazon S3.

S3 Select es compatible con las tablas de Hive basadas en CSV y en archivos JSON, y permite establecer la variable de configuración `s3select.filter` en `true` durante la sesión de Hive. Para obtener más información y ejemplos, consulta Especificación de S3 Select en el código.

¿S3 Select es adecuado para mi aplicación?

Le recomendamos que realice un análisis comparativo sus aplicaciones con y sin S3 Select para ver si su uso puede ser adecuado para su aplicación.

Utilice las siguientes directrices para determinar si la aplicación es candidata a utilizar S3 Select:

- La consulta filtra más de la mitad del conjunto de datos original.
- Los predicados del filtro de consulta utilizan columnas que tienen un tipo de datos compatible con Amazon S3 Select. Para obtener más información, consulte Tipos de datos en la Guía del usuario de Amazon Simple Storage Service.
- La conexión de red entre Amazon S3 y el clúster de Amazon EMR tiene una buena velocidad de transferencia y ancho de banda disponible. Amazon S3 no comprime las respuestas HTTP, por lo que es probable que el tamaño de la respuesta aumente en el caso de los archivos de entrada comprimidos.

Consideraciones y limitaciones

- No se admiten el cifrado del servidor de Amazon S3 con claves de cifrado proporcionadas por el cliente (SSE-C) ni el cifrado del cliente.
- No se admite la propiedad `AllowQuotedRecordDelimiters`. Si se especifica esta propiedad, se producirá un error en la consulta.
- Solo se admiten archivos CSV y JSON en formato UTF-8. No se admiten archivos multilínea ni JSON CSVs.
- Solo se admiten los archivos sin comprimir o gzip o bzip2.
- No se admiten los caracteres de comentario en la última línea.
- Las líneas vacías al final de un archivo no se procesan.
- Hive en Amazon EMR admite los tipos de datos primitivos admitidos por S3 Select. Para obtener más información, consulte Tipos de datos en la Guía del usuario de Amazon Simple Storage Service.

Especificación de S3 Select en el código

Para utilizar S3 Select en la tabla de Hive, cree la tabla especificando `com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat` como nombre de

la clase `INPUTFORMAT` y especifique un valor para la propiedad `s3select.format` mediante la cláusula `TBLPROPERTIES`.

De forma predeterminada, S3 Select se deshabilita cuando se ejecutan consultas. Habilite S3 Select estableciendo `s3select.filter` en `true` en la sesión de Hive, tal y como se muestra a continuación. Los ejemplos que aparecen a continuación muestran cómo especificar S3 Select al crear una tabla a partir de archivos CSV y JSON subyacentes y, a continuación, consultar la tabla con una instrucción select sencilla.

Example Instrucción CREATE TABLE para una tabla basada en CSV

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
STORED AS
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/mycsvfile/'
TBLPROPERTIES (
  "s3select.format" = "csv",
  "s3select.headerInfo" = "ignore"
);
```

Example Instrucción CREATE TABLE para una tabla basada en JSON

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'
STORED AS
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/json/'
```

```
TBLPROPERTIES (
  "s3select.format" = "json"
);
```

Example Instrucción SELECT TABLE

```
SET s3select.filter=true;
SELECT * FROM mys3selecttable WHERE col2 > 10;
```

Optimización de MSCK

Hive almacena una lista de particiones para cada tabla en su metaalmacén. Sin embargo, cuando las particiones se agregan o se eliminan directamente del sistema de archivos, el metaalmacén de Hive no es consciente de estos cambios. El comando MSCK actualiza los metadatos de las particiones en el metaalmacén de Hive para las particiones que se agregaron o eliminaron directamente del sistema de archivos. La sintaxis del comando es:

```
MSCK [REPAIR] TABLE table_name [ADD/DROP/SYNC PARTITIONS];
```

Hive implementa este comando de la siguiente manera:

1. Hive recupera todas las particiones de la tabla del metaalmacén. A partir de la lista de rutas de particiones que no existen en el sistema de archivos, crea una lista de particiones para eliminarlas del metaalmacén.
2. Hive recopila las rutas de las particiones presentes en el sistema de archivos, las compara con la lista de particiones del metaalmacén y genera una lista de las particiones que deben agregarse al metaalmacén.
3. Hive actualiza el metaalmacén mediante el modo `ADD`, `DROP` o `SYNC`.

> **Note**
>
> Cuando hay muchas particiones en el metaalmacén, el paso para comprobar si una partición no existe en el sistema de archivos tarda mucho en ejecutarse, ya que la llamada a la API `exists` del sistema de archivos debe hacerse para cada partición.

En Amazon EMR 6.5.0, Hive introdujo un indicador llamado `hive.emr.optimize.msck.fs.check`. Cuando está activado, este indicador hace que Hive

compruebe la presencia de una partición de la lista de rutas de partición del sistema de archivos que se generó en el paso 2 anterior, en lugar de hacer llamadas a la API del sistema de archivos. En Amazon EMR 6.8.0, Hive tenía esta optimización habilitada de forma predeterminada, lo que eliminaba la necesidad de establecer el indicador `hive.emr.optimize.msck.fs.check`.

Uso de Hive Live Long and Process (LLAP)

Amazon EMR 6.0.0 es compatible con la funcionalidad Live Long and Process (LLAP) para Hive. LLAP utiliza daemons persistentes con almacenamiento inteligente en memoria caché para mejorar el rendimiento de las consultas en comparación con el anterior modo de ejecución predeterminado del contenedor de Tez.

Los daemons de LLAP Hive se administran y ejecutan como un servicio de YARN. Dado que un servicio de YARN puede considerarse una aplicación de YARN de larga duración, algunos de los recursos del clúster están dedicados a Hive LLAP y no se pueden utilizar para otras cargas de trabajo. Para obtener más información, consulte LLAP y la API del servicio YARN.

Habilitación de Hive LLAP en Amazon EMR

Para habilitar Hive LLAP en Amazon EMR, proporcione la siguiente configuración al lanzar un clúster.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true"
    }
  }
]
```

Para obtener más información, consulte Configuración de aplicaciones.

De forma predeterminada, Amazon EMR asigna alrededor del 60 % de los recursos de YARN del clúster a los daemons de Hive LLAP. Puede configurar el porcentaje de recursos YARN de clúster que se asignan a Hive LLAP y el número de nodos principales y de tareas que se tendrán en cuenta para esa asignación a Hive LLAP.

Por ejemplo, la siguiente configuración inicia Hive LLAP con tres daemons en tres nodos de tareas o principales y asigna el 40 por ciento de los recursos de YARN de los tres nodos principales o de tareas a los daemons de Hive LLAP.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true",
      "hive.llap.percent-allocation": "0.4",
      "hive.llap.num-instances": "3"
    }
  }
]
```

Puede utilizar las siguientes configuraciones de `hive-site` en la API de clasificación para anular la configuración predeterminada de recursos de LLAP.

Propiedad	Descripción
hive.llap.daemon.yarn.container.mb	Tamaño total del contenedor de daemons de LLAP (en MB)
hive.llap.daemon.memory.per.instance.mb	La memoria total utilizada por los ejecutores en el contenedor de daemons de LLAP (en MB)
hive.llap.io.memory.size	Tamaño de la caché para entrada/salida de LLAP
hive.llap.daemon.num.executors	Número de ejecutores por daemon de LLAP

Inicio manual de LLAP en el clúster

Todas las dependencias y configuraciones utilizadas por LLAP se empaquetan en el archivo tar de LLAP como parte del inicio del clúster. Si LLAP se ha habilitado mediante `"hive.llap.enabled": "true"`, recomendamos usar la reconfiguración de Amazon EMR para hacer cambios de configuración en LLAP.

De lo contrario, para cualquier cambio manual en `hive-site.xml`, deberá reconstruir el archivo tar de LLAP mediante el comando `hive --service llap`, como se muestra en el ejemplo siguiente.

```
# Define how many resources you want to allocate to Hive LLAP

LLAP_INSTANCES=<how many llap daemons to run on cluster>
LLAP_SIZE=<total container size per llap daemon>
LLAP_EXECUTORS=<number of executors per daemon>
LLAP_XMX=<Memory used by executors>
LLAP_CACHE=<Max cache size for IO allocator>

yarn app -enableFastLaunch

hive --service llap \
--instances $LLAP_INSTANCES \
--size ${LLAP_SIZE}m \
--executors $LLAP_EXECUTORS \
--xmx ${LLAP_XMX}m \
--cache ${LLAP_CACHE}m \
--name llap0 \
--auxhbase=false \
--startImmediately
```

Comprobación del estado de Hive LLAP

Utilice el siguiente comando para comprobar el estado de Hive LLAP mediante Hive.

```
hive --service llapstatus
```

Utilice el siguiente comando para comprobar el estado de Hive LLAP mediante YARN.

```
yarn app -status (name-of-llap-service)

# example:
yarn app -status llap0 | jq
```

Inicio o detención de Hive LLAP

Puesto que Hive LLAP se ejecuta como un servicio de YARN persistente, debe detener o reiniciar el servicio de YARN para detener o reiniciar el Hive LLAP. Los siguientes comandos lo demuestran.

```
yarn app -stop llap0
yarn app -start llap0
```

Cambio del número de daemons de Hive LLAP

Utilice el siguiente comando para reducir el número de instancias de LLAP.

```
yarn app -flex llap0 -component llap -1
```

Para obtener más información, consulte el apartado sobre cómo usar el comando Flex para modificar el número de contenedores de un componente de un servicio.

Cifrado en Hive

En esta sección, se describen los tipos de cifrado que admite Amazon EMR.

Cifrado modular de Parquet en Hive

El cifrado modular de Parquet proporciona control de acceso y cifrado a nivel de columnas para mejorar la privacidad y la integridad de los datos almacenados en el formato de archivo Parquet. Esta característica está disponible en Amazon EMR Hive a partir de la versión 6.6.0.

Las soluciones de seguridad e integridad compatibles anteriormente, que incluyen el cifrado de archivos o el cifrado de la capa de almacenamiento, se describen en Opciones de cifrado de la Guía de administración de Amazon EMR. Estas soluciones se pueden utilizar para los archivos de Parquet, pero al aprovechar las nuevas características del mecanismo de cifrado integrado de Parquet, se obtiene un acceso detallado a nivel de columnas, así como mejoras en el rendimiento y la seguridad. Obtenga más información sobre esta característica en la página de GitHub de Apache, Cifrado modular de Parquet.

Los usuarios transfieren las configuraciones a los lectores y escritores Parquet mediante las configuraciones de Hadoop. Las configuraciones detalladas para que los usuarios puedan configurar lectores y escritores para habilitar el cifrado, así como activar funciones avanzadas, se encuentran documentadas en PARQUET-1854: interfaz basada en propiedades para la administración de cifrado de Parquet

Ejemplos de uso

El siguiente ejemplo describe la creación y escritura en una tabla de Hive que utiliza AWS KMS para administrar claves de cifrado.

1. Implemente un KmsClient para el AWS KMS servicio tal como se describe en el documento PARQUET-1373: Herramientas de administración de claves de cifrado. En el siguiente ejemplo, se muestra un fragmento de implementación.

```
package org.apache.parquet.crypto.keytools;

import com.amazonaws.AmazonClientException;
import com.amazonaws.AmazonServiceException;
import com.amazonaws.regions.Regions;
import com.amazonaws.services.kms.AWSKMS;
import com.amazonaws.services.kms.AWSKMSClientBuilder;
import com.amazonaws.services.kms.model.DecryptRequest;
import com.amazonaws.services.kms.model.EncryptRequest;
import com.amazonaws.util.Base64;
import org.apache.hadoop.conf.Configuration;
import org.apache.parquet.crypto.KeyAccessDeniedException;
import org.apache.parquet.crypto.ParquetCryptoRuntimeException;
import org.apache.parquet.crypto.keytools.KmsClient;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;

import java.nio.ByteBuffer;
import java.nio.charset.Charset;
import java.nio.charset.StandardCharsets;

public class AwsKmsClient implements KmsClient {

    private static final AWSKMS AWSKMS_CLIENT = AWSKMSClientBuilder
            .standard()
            .withRegion(Regions.US_WEST_2)
            .build();
    public static final Logger LOG = LoggerFactory.getLogger(AwsKmsClient.class);

    private String kmsToken;
    private Configuration hadoopConfiguration;

    @Override
```

```
   public void initialize(Configuration configuration, String kmsInstanceID, String
kmsInstanceURL, String accessToken) throws KeyAccessDeniedException {
       hadoopConfiguration = configuration;
       kmsToken = accessToken;

   }

   @Override
   public String wrapKey(byte[] keyBytes, String masterKeyIdentifier) throws
KeyAccessDeniedException {
       String value = null;
       try {
           ByteBuffer plaintext = ByteBuffer.wrap(keyBytes);

           EncryptRequest req = new
EncryptRequest().withKeyId(masterKeyIdentifier).withPlaintext(plaintext);
           ByteBuffer ciphertext = AWSKMS_CLIENT.encrypt(req).getCiphertextBlob();

           byte[] base64EncodedValue = Base64.encode(ciphertext.array());
           value = new String(base64EncodedValue, Charset.forName("UTF-8"));
       } catch (AmazonClientException ae) {
           throw new KeyAccessDeniedException(ae.getMessage());
       }
       return value;
   }

   @Override
   public byte[] unwrapKey(String wrappedKey, String masterKeyIdentifier) throws
KeyAccessDeniedException {
       byte[] arr = null;
       try {
           ByteBuffer ciphertext  =
ByteBuffer.wrap(Base64.decode(wrappedKey.getBytes(StandardCharsets.UTF_8)));
           DecryptRequest request = new
DecryptRequest().withKeyId(masterKeyIdentifier).withCiphertextBlob(ciphertext);
           ByteBuffer decipheredtext =
AWSKMS_CLIENT.decrypt(request).getPlaintext();
           arr = new byte[decipheredtext.remaining()];
           decipheredtext.get(arr);
       } catch (AmazonClientException ae) {
           throw new KeyAccessDeniedException(ae.getMessage());
       }
       return arr;
   }
```

```
}
```

2. Cree las claves de AWS KMS cifrado para el pie de página y las columnas a las que tengan acceso sus funciones de IAM, tal y como se describe en la sección Creación de claves de la Guía para desarrolladores.AWS Key Management Service El rol de IAM predeterminado es EMR_ECS_Default.
3. En la aplicación Hive de un clúster de Amazon EMR, agregue el cliente anterior mediante la instrucción `ADD JAR`, tal y como se describe en la documentación de recursos de Apache Hive. A continuación se muestra un ejemplo de instrucción.

   ```
   ADD JAR 's3://location-to-custom-jar';
   ```

   Un método alternativo consiste en agregar el JAR al `auxlib` de Hive mediante una acción de arranque. La siguiente es una línea de ejemplo que se agregará a la acción de arranque:

   ```
   aws s3 cp 's3://location-to-custom-jar' /usr/lib/hive/auxlib
   ```

4. Ajuste los siguientes valores de configuración:

   ```
   set
    parquet.crypto.factory.class=org.apache.parquet.crypto.keytools.PropertiesDrivenCryptoFacto
   set
    parquet.encryption.kms.client.class=org.apache.parquet.crypto.keytools.AwsKmsClient;
   ```

5. Cree una tabla Hive con formato Parquet y especifique las AWS KMS claves en SERDEPROPERTIES e inserte algunos datos en ella:

   ```
   CREATE TABLE my_table(name STRING, credit_card STRING)
   ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
   WITH SERDEPROPERTIES (
     'parquet.encryption.column.key’=<aws-kms-key-id-for-column-1>: credit_card’,
     'parquet.encryption.footer.key’='<aws-kms-key-id-for-footer>’)
   STORED AS parquet
   LOCATION “s3://<bucket/<warehouse-location>/my_table”;

   INSERT INTO my_table SELECT
   java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',5) as
    name,
   java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',10) as
    credit_card
   ```

```
from (select 1) x lateral view posexplode(split(space(100),' ')) pe as i,x;

select * from my_table;
```

6. Compruebe que al crear una tabla externa en la misma ubicación sin acceso a AWS KMS las claves (por ejemplo, si se deniega el acceso al rol de IAM), no pueda leer los datos.

```
CREATE EXTERNAL TABLE ext_table (name STRING, credit_card STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
STORED AS parquet
LOCATION “s3://<bucket>/<warehouse-location>/my_table”;

SELECT * FROM ext_table;
```

7. La última instrucción debería generar la siguiente excepción:

```
Failed with exception
 java.io.IOException:org.apache.parquet.crypto.KeyAccessDeniedException: Footer key:
 access denied
```

Cifrado en tránsito en 2 HiveServer

A partir de la versión 6.9.0 de Amazon EMR, HiveServer 2 (HS2) está habilitado para TLS/SSL como parte de la configuración de seguridad. Cifrado en tránsito en 2 HiveServer Esto afecta a la forma en que HS2 se conecta a un clúster de Amazon EMR con el cifrado en tránsito activado. Para conectarse HS2, debe modificar los valores `TRUSTSTORE_PATH` y los `TRUSTSTORE_PASSWORD` parámetros de la URL del JDBC. La siguiente URL es un ejemplo de una conexión JDBC HS2 con los parámetros necesarios:

```
jdbc:hive2://HOST_NAME:10000/
default;ssl=true;sslTrustStore=TRUSTSTORE_PATH;trustStorePassword=TRUSTSTORE_PASSWORD
```

Utilice las instrucciones correspondientes para el cifrado 2D dentro o fuera del clúster, que se indican a continuación. HiveServer

On-cluster HS2 access

Si va a acceder a HiveServer 2 mediante el cliente Beeline después de utilizar SSH con el nodo principal, consulte para buscar los valores de los parámetros `/etc/hadoop/conf/ssl-`

`server.xml` y mediante la configuración `TRUSTSTORE_PATH` y`TRUSTSTORE_PASSWORD`. `ssl.server.truststore.location` `ssl.server.truststore.password`

Los siguientes comandos de ejemplo pueden ayudarle a recuperar estas configuraciones:

```
TRUSTSTORE_PATH=$(sed -n '/ssl.server.truststore.location/,+2p' /etc/hadoop/conf/
ssl-server.xml | awk -F "[><]" '/value/{print $3}')
TRUSTSTORE_PASSWORD=$(sed -n '/ssl.server.truststore.password/,+2p' /etc/hadoop/
conf/ssl-server.xml | awk -F "[><]" '/value/{print $3}')
```

Off-cluster HS2 access

Si está accediendo a HiveServer 2 desde un cliente externo al clúster de Amazon EMR, puede utilizar uno de los siguientes enfoques para obtener el y: `TRUSTSTORE_PATH` `TRUSTSTORE_PASSWORD`

- Convierta el archivo PEM que se creó durante la configuración de seguridad en un archivo JKS y utilícelo en la URL de conexión JDBC. Por ejemplo, con openssl y keytool, use los siguientes comandos:

  ```
  openssl pkcs12 -export -in trustedCertificates.pem -inkey privateKey.pem -out
   trustedCertificates.p12 -name "certificate"
  keytool -importkeystore -srckeystore trustedCertificates.p12 -srcstoretype pkcs12
   -destkeystore trustedCertificates.jks
  ```

- Como alternativa, consulte `/etc/hadoop/conf/ssl-server.xml` para buscar los valores de los parámetros `TRUSTSTORE_PATH` y `TRUSTSTORE_PASSWORD` mediante la configuración `ssl.server.truststore.location` y `ssl.server.truststore.password`. Descargue el archivo truststore en la máquina cliente y utilice la ruta de la máquina cliente como `TRUSTSTORE_PATH`.

  Para obtener más información sobre el acceso a las aplicaciones desde un cliente fuera del clúster de Amazon EMR, consulte Uso del controlador JDBC de Hive.

Historial de versiones de Hive

En la siguiente tabla se muestra la versión de Hive incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión

correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Hive

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
		server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-5.36.2	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.36.1	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.36.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.35.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.3.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.2.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Etiqueta de versión de Amazon EMR	Versión de Hive	Componentes instalados con Hive
emr-4.1.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.0.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Notas de la versión de Hive por versión

Temas

- Amazon EMR 7.8.0 - Notas de la versión de Hive
- Amazon EMR 7.7.0 - Notas de la versión de Hive
- Amazon EMR 7.6.0 - Notas de la versión de Hive
- Amazon EMR 7.5.0 - Notas de la versión de Hive
- Amazon EMR 7.4.0 - Notas de la versión de Hive
- Amazon EMR 7.3.0 - Notas de la versión de Hive

- Amazon EMR 7.2.0 - Notas de la versión de Hive
- Amazon EMR 7.1.0: notas de la versión de Hive
- Amazon EMR 7.0.0 - Notas de la versión de Hive
- Amazon EMR 6.15.0: notas de la versión de Hive
- Amazon EMR 6.14.0: notas de la versión de Hive
- Amazon EMR 6.13.0: notas de la versión de Hive
- Amazon EMR 6.12.0: notas de la versión de Hive
- Amazon EMR 6.11.0: notas de la versión de Hive
- Amazon EMR 6.10.0: notas de la versión de Hive
- Amazon EMR 6.9.0: notas de la versión de Hive
- Amazon EMR 6.8.0: notas de la versión de Hive
- Amazon EMR 6.7.0: notas de la versión de Hive
- Amazon EMR 6.6.0: notas de la versión de Hive

Amazon EMR 7.8.0 - Notas de la versión de Hive

Amazon EMR 7.8.0: cambios en la colmena

Tipo	Descripción
Corrección de errores	Corrige el error CVE-2024-23953: Apache Hive: Timing Attack Against Signature in LLAP.

Problemas conocidos

- Para las consultas de sobreescritura de Hive Insert con Amazon S3 Express One Zone como ubicación de salida, defina la configuración del sitio principal: en. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.7.0 - Notas de la versión de Hive

Amazon EMR 7.7.0: cambios en la colmena

Tipo	Descripción
Corrección de errores	Corrige el error CVE-2024-29869: Apache Hive: archivo de credenciales creado con permisos no restrictivos.
Corrección de errores	Se corrige SemanticException cuando se habilita una política de filtrado a nivel de fila en Apache Ran.
Corrección de errores	Deshabilite Tez Async Init RR cuando LLAP o ACID estén habilitados.

Problemas conocidos

- Para las consultas de sobreescritura de Hive Insert con Amazon S3 Express One Zone como ubicación de salida, defina la configuración del sitio principal: en. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.6.0 - Notas de la versión de Hive

Amazon EMR 7.6.0: cambios en la colmena

Tipo	Descripción
Mejoras	Se agregó una función rápida de listado de prefijos S3 para las tablas particionadas ORC que no son de ACID
Característica	Se agregó compatibilidad con Magic Committer s para consultas de escritura en Hive en el sistema S3 AFile

Problemas conocidos

- Para las consultas de sobreescritura de Hive Insert con Amazon S3 Express One Zone como ubicación de salida, defina la configuración del sitio principal: en. `fs.s3a.directory.operations.purge.uploads` `false`

Amazon EMR 7.6.0: nuevas configuraciones

Clasificación	Nombre	Predeterminado/a	Descripción
hive-site	`hive.exec.fast.s3.partition.discovery.enabled`	true	Si se debe utilizar la detección rápida de particiones de S3 para el cálculo de divisiones. Esto habilitará la lista basada en prefijos para los formatos de archivo compatibles: ORC. Tenga en cuenta que esta función utiliza un parámetro de la API de S3 que la clase de almacenamiento S3 Express One Zone no admite. Cuando los utilices, desactiva esta función.
hive-site	`hive.exec.fast.s3.partition.discovery.max.thread.threshold`	128	El grado máximo de paralelismo para una rápida detección de particiones S3.

Clasificación	Nombre	Predeterminado/a	Descripción
hive-site	`hive.exec.fast.s3.partition.discovery.parallelism`	10	El grado de paralelismo de una sola ejecución de detección rápida de particiones S3. Esta configuración solo tiene efecto si `hive.exec.fast.s3.partition.discovery.enabled` se establece en `true`

Clasificación	Nombre	Predeterminado/a	Descripción
hive-site	`hive.blobstore.output-committer.magic.track.commits.in.memory.enabled`	true	Si Magic Committer y Hive rastrean todas las confirmaciones pendientes en la memoria, ¿deberían registrarse? El confirmador de Magic tiene la opción de almacenar los datos de confirmación en la memoria, lo que puede acelerar la TaskCommit operación al realizar menos llamadas a S3. Esta configuración anula la configuración de Hadoop `fs.s3a.committer.magic.track.commits.in.memory.enabled`

Clasificación	Nombre	Predeterminado/a	Descripción
hive-site	`hive.blobstore.output-committer.dp.skip.task.staging.dir.creation`	true	¿Debería Magic Committer crear las rutas de ensayo dp en la blobstore? Esta marca solo se aplica al rastrear las confirmaciones en la memoria cuando Hive usa Magic Committer a través de. `hive.blobstore.output-committer.magic.track.commits.in.memory.enabled` De forma predeterminada, se establece en true, pero solo surte efecto si `hive.blobstore.output-committer.magic.track.commits.in.memory.enabled` está activado y guarda las llamadas S3 adicionales de las rutas de intentos de creación de tareas en blobstore.

Clasificación	Nombre	Predeterminado/a	Descripción
hive-site	`hive.blobstore.output-committer.magic.disable.fs.cache.for.llap`	true	Marca esta opción para seleccionar si las cachés FS de blobstore deben estar deshabilitadas en los flujos de escritura de LLAP cuando se usa Magic Committer. Este indicador aparece cuando el LLAP está activado y, por defecto, está establecido en true.

Amazon EMR 7.5.0 - Notas de la versión de Hive

Amazon EMR 7.5.0: cambios en la colmena

Tipo	Descripción
Mejoras	Se aumentó el tiempo máximo de espera a que se abra la sesión de Tez al intentar utilizar la sesión existente en HiveCLI a 10 segundos
Mejora	Se han ajustado las configuraciones para mejorar el rendimiento en consultas de selección sencillas con LIMIT

Amazon EMR 7.4.0 - Notas de la versión de Hive

Amazon EMR 7.4.0: Cambios en la colmena

Tipo	Descripción
Upgrade	HIVE-28191: actualice la versión de Hadoop a la 3.4.0
Upgrade	Actualice el protobug compartido de Hadoop a la versión 3.21
Upgrade	Actualice commons-cli a la versión 1.5.0
Upgrade	Actualice commons-compress a 1.24.0
Upgrade	Actualice commons-io a 2.14.0
Upgrade	Actualice commons-lang3 a la versión 3.21.0
Mejora	Cambie el tiempo de espera a que se abra la sesión de Tez al intentar usar la sesión existente en HiveCLI a 10 segundos
Mejora	Habilite el mecanismo de cortocircuito en Tez DAG para consultas de selección sencillas con LIMIT
Mejora	HIVE-21100: Se permite aplanar los subdirect orios de las tablas al utilizar el motor TEZ y la cláusula UNION
Corrección de errores	HIVE-25095: El comando Beeline/hive -e no puede procesar consultas con comillas al final
Corrección de errores	HIVE-13781: Tez Job falló FileNotFoundExcept ion cuando el directorio de particiones no existe

Tipo	Descripción
Corrección de errores	HIVE-28480: Deshabilita SMB si el generador de hash de partición no coincide entre las ramas de unión en el RS anterior

Amazon EMR 7.4.0: Nuevas configuraciones

Clasificación	Nombre	Predeterminado/a	Descripción
hive-site	hive.ignore.failure.partition.dir.not.found	false	Ignora el error si la partición de la tabla existe pero la ruta de almacenamiento de objetos real no existe.
hive-site	hive.tez.union.flatten.subdirectorios	false	Al escribir datos en una tabla y UNION ALL es el último paso de la consulta, Hive on Tez crea un subdirectorio para cada rama de UNION ALL. Cuando esta propiedad está habilitada, se eliminan los subdirectorios y se cambia el nombre de los archivos y se mueven al directorio principal. Tenga en cuenta que esto no tiene ningún efecto cuando hive.blobstore.use.output-

Clasificación	Nombre	Predeterminado/a	Descripción
			committer está activado.

Amazon EMR 7.3.0 - Notas de la versión de Hive

Amazon EMR 7.3.0: cambios en la colmena

Tipo	Descripción
Característica	HIVE-18728: web segura con SSL. HCat
Mejora	Support configure las credenciales del almacén de claves SSL para la interfaz de usuario web del daemon LLAP.
Mejora	Ofrece la opción de controlar la verificación del nombre de host SSL para el servidor Hive Metastore.
Corrección de errores	HIVE-26541: NPE al iniciar el servicio web. HCat
Corrección de errores	HIVE-23011 — El optimizador de trabajo compartido debería comprobar los predicados residuales al comparar las uniones.
Corrección de errores	Corrija javax.security.sasl. SaslException: No hay una capa de protección común entre el cliente y el servidor entre el HMS y Namenode cuando el cifrado en tránsito está activado.
Corrección de errores	Se corrige el problema por el IOExceptionque el final de la división de los orcos se superpone con el inicio de la ubicación de un bloque.
Corrección de errores	Utilice el delimitador de nombres de columnas en lugar de dividirlos siempre por comas

Tipo	Descripción
	cuando los nombres de las columnas contengan un carácter de coma y utilice. CSVSerde

Amazon EMR 7.3.0: nuevas configuraciones

Clasificación	Nombre	Predeterminado/a	Descripción
hcatalog-webhcat-site	templeton.use.ssl	false	Establézcalo en true para usar el cifrado SSL para el servidor web. HCat
hcatalog-webhcat-site	templeton.keystore .path		Ubicación del almacén de claves del certificado SSL para el servidor web. HCat
hcatalog-webhcat-site	templeton.keystore .password		Contraseña del almacén de claves del certificado SSL para el servidor web. HCat
hcatalog-webhcat-site	templeton.ssl.prot ocol.blacklist	SSLv2, SSLv3	Versiones de SSL para deshabilitar en el servidor web. HCat
hcatalog-webhcat-site	templeton.host	0.0.0.0	La dirección de host en la que escuchará el HCat servidor web.
hive-site	hive.metastore.ssl .enable.hostname.v erification	false	Controle la verificac ión del nombre de host durante el

Clasificación	Nombre	Predeterminado/a	Descripción
			apretón de manos SSL/TLS.
hive-site	hive.llap.daemon.web.ssl.keystore.path		Ubicación del almacén de claves del certificado SSL para la interfaz de usuario web del daemon de LLAP.
hive-site	hive.llap.daemon.web.ssl.keystore.password		Contraseña del almacén de claves del certificado SSL para la interfaz de usuario web del daemon de LLAP.
hive-site	hive.metastore.hadoop.rpc.protection.override.to.authentication	false	Cuando está activado, HMS siempre anula el valor de hadoop.rpc.protection para la autenticación en su conjunto de configuraciones.

Amazon EMR 7.2.0 - Notas de la versión de Hive

Amazon EMR 7.2.0: cambios en la colmena

Tipo	Descripción
Upgrade	Parquet 1.13.1: Parquet se ha actualizado a la versión 113.1.
Mejora	HIVE-12930: Support SSL shuffle para LLAP.

Tipo	Descripción
Mejora	HIVE-23062: Hive para comprobar la URL de Yarn RM en modo TLS y Yarn HA para crear una cola de Tez personalizada.
Corrección de errores	HIVE-27952: Hive no se crea cuando tiene varios certificados. SslContextFactory KeyStore
Corrección de errores	HIVE-28085: YarnQueueHelper no puede acceder al YARN habilitado para HTTPS. WebService
Corrección de errores	HIVE-26436: Hive en MR NullPointerExcepti on cuando no se ha llamado a InitializeOp y se cierra la llamada. Si el operador no se ha inicializado, omita el botón de cierre..

Amazon EMR 7.2.0: nuevas configuraciones

Clasificación	Nombre	Predeterminado/a	Descripción
hive-site	hive.llap.shuffle. ssl.habilitado	false	Configúrelo en true junto con tez.runti me.shuffle.ssl.ena ble para habilitar la mezcla de SSL para LLAP.

Amazon EMR 7.1.0: notas de la versión de Hive

Amazon EMR 7.1.0: cambios en Hive

Tipo	Descripción
Corrección de errores	HIVE-24381 — La entrada de texto comprimid o devuelve 0 filas si se incluye el encabezado o pie de página para omitir.
Corrección de errores	ShuffleHandler HIVE-24190 — LLAP: podría devolver DISK_ERROR_EXCEPTION según TEZ-4233.
Corrección de errores	HIVE-23073 — Sombrear Netty.
Corrección de errores	HIVE-23073 — Sombrear Netty y actualizar a Netty 4.1.48.Final.
Corrección de errores	HIVE-23148 — El flujo de clientes externos de Llap se interrumpe debido al sombreado de netty.
Corrección de errores	HIVE-25180 — Actualizar Netty.
Corrección de errores	HIVE-24524 — LLAP: actualice a Netty4 y elimine la dependencia de Netty3 de la colmena siempre que sea posible. ShuffleHa ndler
Corrección de errores	HIVE-28000 — Hive QL: la cláusula “no está dentro” arroja resultados incorrectos cuando no se puede aplicar coerción de tipos.
Corrección de errores	HIVE-27993 ShuffleHandler — Netty4 debería usar 1 hilo principal.
Upgrade	Actualiza Netty a la versión 4.1.100.Final
Upgrade	Actualización de Jetty a 9.4.53.v20231009

Tipo	Descripción
Upgrade	Se ha actualizado ZooKeeper a la versión 3.9.1

Amazon EMR 7.1.0: cambios en Hive

- Amazon EMR 7.1 actualizar Hive a Netty 4.1.100.Final para solucionar las vulnerabilidades de seguridad de Netty3. Como hive-druid-handler depende de netty3, Hive no tiene el `hive-druid-handler` JAR en la ruta de clases de Hive en Amazon EMR 7.1. Una próxima versión de Amazon EMR lo incluirá en la ruta de clases de Hive una vez que el controlador Druid sea compatible con la versión 4.1.100.Final o posteriores de Netty. Póngase en contacto con el servicio de AWS asistencia si necesita el `hive-druid-handler` JAR en Amazon EMR, versión 7.1 o superior.

Amazon EMR 7.0.0 - Notas de la versión de Hive

Amazon EMR 7.0.0: cambios en la colmena

Tipo	Descripción
Upgrade	Hive Runtime ahora usa Java 17 de forma predeterminada. Consulte la Guía de versión de EMR 7.0.0 para obtener más información.
Portabilidad con versiones anteriores	HIVE-17709: elimine las referencias de Sun.Misc.Cleaner
Corrección de errores	Desactive Tez Async Init RR cuando LLAP o ACID estén activados

Amazon EMR 6.15.0: notas de la versión de Hive

Amazon EMR 6.15.0: cambios en Hive

Tipo	Descripción
Característica	Soporte para TEZ-4397: para la apertura de división asíncrona de Tez, Hive ahora admite

Tipo	Descripción
	las configuraciones de Tez que se describen en Apertura de división asíncrona de Tez.
Corrección de errores	HIVE-25400: mueva el desfase de actualización en `BytesColumnVector` a `setValPreallocated` .
Corrección de errores	HIVE-25190: corrija muchas asignaciones pequeñas en `BytesColumnVector` .
Corrección de errores	Empaquetar los módulos netty con el servidor llap para evitar NoClassDefFoundexcepciones al iniciar LLapDaemon en los nodos de trabajo.
Upgrade	Se ha actualizado Hadoop a la versión 3.3.6.
Upgrade	HIVE-26684: actualización de `maven-shade-plugin` a 3.4.1.
Mejora	Para reducir el tiempo de inicio del clúster de Amazon EMR, elimine 15 segundos de tiempo de reposo del HCatalog script de inicio.

Amazon EMR 6.14.0: notas de la versión de Hive

Amazon EMR 6.14.0: cambios en Hive

Tipo	Descripción
Mejoras	HIVE-26762: eliminación del operando, arreglos en `HiveFilterSetOpTransposeRule`
Corrección de errores	HIVE-27582: No almacene en caché el formato de entrada de la tabla HBase FetchOperator

Tipo	Descripción
Corrección de errores	HIVE-26452: NPE al convertir JOIN en MAPJOIN y la columna JOIN referenciadas más de una vez
Corrección de errores	HIVE-26416: `AcidUtils.isRawFormatFile()` genera `InvalidProtocolBufferException` para el archivo non-ORC
Corrección de errores	HIVE-26105: Mostrar columnas muestra valores adicionales si la columna Comentarios contiene un carácter chino específico
Corrección de errores	HIVE-25864: optimización de las consultas de Hive que crea un plan incorrecto para la inserción de predicados con la función de ventana
Corrección de errores	HIVE-25224: múltiples declaraciones INSERT que incluyen tablas con diferentes `bucketing_versions` generan errores
Corrección de errores	HIVE-24151: `MultiDelimitSerDe` cambia los datos si las cadenas contienen caracteres non-ASCII
Corrección de errores	HIVE-23606: (LLAP) retraso en la limpieza de `DirectByteBuffer` para `EncodedReaderImpl`
Corrección de errores	HIVE-22165: la sincronización introducida por HIVE-14296 en `SessionManager.closeSession` provoca una alta latencia en un servidor de Hive ocupado

Tipo	Descripción
Corrección de errores	HIVE-21304: fortalecimiento del uso de la versión de asignación de buckets

Amazon EMR 6.13.0: notas de la versión de Hive

Amazon EMR 6.13.0: cambios en Hive

Tipo	Descripción
Mejoras	Actualización de los scripts de Python para que sean compatibles con Python3
Mejora	HIVE-27097: Mejore la estrategia de reintentos para el cliente y el servidor MetaStore
Corrección de errores	HIVE-21778: CBO: “La estructura no es nula” se evalúa como anulable, lo que siempre provoca que el filtro no aparezca en la consulta
Corrección de errores	HIVE-21009: se agrega la posibilidad de que el usuario defina el usuario de enlace
Corrección de errores	HIVE-22661: la compactación falla en una tabla no agrupada con datos cargados en la ruta
Corrección de errores	HIVE-19718: al agregar particiones de forma masiva, también se obtiene una tabla para cada partición
Corrección de errores	HIVE-22173: la consulta con múltiples vistas laterales se bloquea durante la compilación
Corrección de errores	HIVE-27088: los resultados son incorrectos cuando se combinan las uniones internas y externas con filtros posteriores a la unión

Tipo	Descripción
Corrección de errores	HIVE-21935: vectorización de Hive: rendimien to degradado con la UDF vectorizada
Corrección de errores	HIVE-25299: la conversión de la marca de tiempo a tipos de datos numéricos es incorrecta en zonas horarias distintas de UTC
Corrección de errores	HIVE-24626: LLAP: los subprocesos del lector podrían quedarse inactivos si todos los subprocesos del elevador de E/S están ocupados para ponerse en cola con otros lectores con la cola llena
Corrección de errores	HIVE-27029: la consulta de Hive falla y el sistema de archivos está cerrado. Se ha hecho una revisión para el HIVE-26352
Corrección de errores	HIVE-26352: la comprobación de acceso a la cola de Tez falla y se produce una excepción de GSS en caso de compactación
Corrección de errores	HIVE-24590: el registro de operaciones sigue filtrando los appenders log4j
Corrección de errores	HIVE-24552: Posibles fugas o acumulaciones en las conexiones del HMS loadDynamicPartiti ons
Corrección de errores	HIVE-27069: al unir el mapa de buckets, se obtienen resultados incorrectos
Corrección de errores	HIVE-27344: Añada una marca nula en #close RecordReaderImpl
Corrección de errores	HIVE-27439: compatibilidad con espacio en decimal

Tipo	Descripción
Corrección de errores	HIVE-27267: se obtienen resultados incorrectos al unir el mapa de buckets en una columna agrupada en decimales con una subconsulta
Corrección de errores	HIVE-21986: Interfaz de usuario HiveServer web: configuración del encabezado de respuesta predeterminado Strict-Transport-Security
Corrección de errores	HIVE-22148: los tokens de delegación S3A no se agregan a la configuración de tareas del compactador.
Corrección de errores	HIVE-22622: Hive permite crear una estructura con nombres de atributos duplicados
Corrección de errores	HIVE-22008: el operador LIKE debe coincidir con la entrada multilínea
Corrección de errores	HIVE-23144: LLAP: Let cleanup en ServiceStop QueryTracker
Corrección de errores	HIVE-22391: NPE al comprobar la caché de resultados de consultas de Hive
Corrección de errores	HIVE-23305: en AddNode debido a las condiciones de carrera NullPointerException LlapTaskSchedulerService
Corrección de errores	HIVE-22178: El parquet lo persigue FilterPredicate CastException SchemaEvolution
Corrección de errores	HIVE-21517: Corregido AggregateStatsCache
Corrección de errores	HIVE-21825: mejora del mensaje de error del cliente cuando la HA activa/pasiva está habilitada

Tipo	Descripción
Corrección de errores	HIVE-23389: puede provocar FilterMergeRule AssertionError
Corrección de errores	HIVE-22767: Beeline no analiza correctamente los puntos y comas de los comentarios
Corrección de errores	HIVE-22996: el BasicStats análisis debería comprobar de forma proactiva si hay cadenas nulas o vacías
Corrección de errores	HIVE-22808: no maneja HiveRelFieldTrimmer HiveTableFunctionScan
Corrección de errores	HIVE-22437: los metadatos LLAP almacenan en caché el NPE al bloquear los metadatos.
Corrección de errores	HIVE-22606: registra avro.schema.literal en el nivel INFO AvroSerde
Corrección de errores	HIVE-22713: no se debe hacer una propagación constante para la estructura Join-Fil(*)-RS
Corrección de errores	HIVE-21624: LLAP: las métricas de la CPU a nivel de subproceso no funcionan
Corrección de errores	HIVE-22815: reduce la creación innecesaria de objetos del sistema de archivos en MROutput
Corrección de errores	HIVE-23060: fallo en la consulta con el error “La expresión de agrupación de conjuntos no está en la clave GROUP BY. Se encontró el error cerca del token”.
Corrección de errores	HIVE-22236: no se pudo crear la vista al seleccionar la vista que contiene la subconsulta `NOT IN`

Tipo	Descripción
Corrección de errores	HIVE-19886: los registros pueden dirigirse a 2 archivos si se utiliza —hiveconf hive.log.file
Corrección de errores	HIVE-20620: se manifiestan colisiones al insertar en tablas MM ordenadas por buckets con particiones dinámicas
Corrección de errores	HIVE-14557: puntero nulo cuando ambos y Mapjoin están activados SkewJoin
Corrección de errores	HIVE-20471: problemas para obtener la ruta de base de datos predeterminada
Corrección de errores	HIVE-20598: Corrige errores tipográficos en los cálculos HiveAlgorithmsUtil
Corrección de errores	HIVE-14737: problema al acceder a /logs en una interfaz de usuario web Kerberizada de Hive Server 2
Corrección de errores	HIVE-20733: No se puede utilizar el código UDFOPEqual NS genérico = en las descripci ones de los planes
Corrección de errores	HIVE-20848: Tras configurar la UpdateInp utAccessTimeHook consulta, se produce un error y no se encuentra la tabla.
Corrección de errores	HIVE-18929: El método de .java tiene una condición de carrera humanReadableInt . HiveStringUtils
Corrección de errores	HIVE-20841: LLAP: permite configurar los puertos dinámicos
Corrección de errores	HIVE-20930: en modo FILTRO no surte efecto VectorCoalesce

Tipo	Descripción
Corrección de errores	HIVE-21007: una unión parcial puede llevar a planes equivocados
Corrección de errores	HIVE-21074: la eliminación de consultas en tablas agrupadas en Hive no funciona para la condición IS NOT NULL
Corrección de errores	HIVE-21223: CachedStore devuelve una partición nula cuando la partición no existe
Corrección de errores	HIVE-19625: posible NPE y ocultación de una excepción real en Hive#copyFiles
Corrección de errores	HIVE-17020: una deduplicación agresiva del RS puede eliminar incorrectamente la rama de un árbol OP
Corrección de errores	HIVE-20168: Registro oculto ReduceSinkOperator
Corrección de errores	HIVE-20879: El uso de un valor nulo en una expresión de proyección conduce a CastException
Corrección de errores	HIVE-20888 TxnHandler: se invoca sort () en listas inmutables
Corrección de errores	HIVE-19948: no HiveCli divide correctamente el comando por punto y coma si las comillas están dentro de la cadena
Corrección de errores	HIVE-20621: se llama a resultset.next, lo que provoca una lentitud incremental GetOperationStatus

Tipo	Descripción
Corrección de errores	HIVE-20854: valores predeterminados razonables: el intervalo de latidos de Hive en Zookeeper es de 20 minutos, cambia a 2
Corrección de errores	HIVE-20330: El HCat cargador no puede procesar varios objetos para un trabajo con múltiples entradas InputJobInfo
Corrección de errores	HIVE-20787: MapJoinBytesTableContainer El estuche DummyRow no admite la reutilización
Corrección de errores	HIVE-20331: la consulta con la combinación de todos, la vista lateral y la combinación fallan porque “no se puede encontrar el elemento principal en el operador secundario”
Corrección de errores	HIVE-19968: la excepción a la UDF no se lanza
Corrección de errores	HIVE-20410: si se anula la operación de inserción y sobrescritura en la tabla de transacciones, se produce el error “No hay suficiente historial disponible para...”
Corrección de errores	HIVE-20059: la transmisión de Hive debería probar el prefijo de sombreado incondici onalmente en caso de excepción
Corrección de errores	HIVE-19424: Entrada NPE MetaDataF ormatters
Corrección de errores	HIVE-20355: parámetro de limpieza de .setSchema HiveConnection
Corrección de errores	HIVE-20858: El serializador no se inicializó correctamente con la configuración en Utilities. createEmptyBuckets

Tipo	Descripción
Corrección de errores	HIVE-20424: schematool no debe contaminar el historial de líneas rectas
Corrección de errores	HIVE-20338: LLAP: se fuerza el identificador de fichero sintético para sistemas de archivos que tienen impls del protocolo HDFS con semántica de mutación POSIX
Corrección de errores	HIVE-11708: Los operadores lógicos aumentan con NULL ClassCastExceptions
Corrección de errores	HIVE-21082: en HPL/SQL, la instrucción declare no admite variables de tipo carácter
Corrección de errores	HIVE-16690: configuración de la periferia del producto cartesiano Tez en función del tamaño del clúster de LLAP
Corrección de errores	HIVE-14516:. OrcInputFormat SplitGenerator. Llame a Internal
Corrección de errores	HIVE-20981: transmisión/filtraciones AbstractR ecordWriter HeapMemoryMonitor
Corrección de errores	HIVE-20043: HiveServer 2: tiene un bloque de sincronización estática alrededor de un SessionState AtomicBoolean
Corrección de errores	HIVE-20191: la aplicación del parche no falla si el PreCommit parche está vacío
Corrección de errores	HIVE-20400: la tabla de creación siempre debe utilizar una ruta totalmente cualificada para evitar una posible ambigüedad de FS

Tipo	Descripción
Corrección de errores	Adición de una comprobación nula de skewedInfo antes de acceder a las columnas asimétricas

Amazon EMR 6.12.0: notas de la versión de Hive

Amazon EMR 6.12.0: cambios en Hive

Tipo	Descripción
Mejoras	Se ha agregado compatibilidad con JDK 11 y JDK 17 Runtime
Mejoras	Agregue compatibilidad con consultas de nombres de columnas de palabras clave reservadas y que distingan mayúscula s de minúsculas al usar S3 Select. Para usarlo, defina la propiedad de la tabla en el formato “s3select.column.m apping” = "*column1*:*fieldName1* , *column2*:*fieldName2* ,..."
Mejoras	HIVE-23133: las operaciones numéricas pueden tener resultados diferentes en todos los arcos de hardware
Mejora	HIVE-27145: Se utiliza StrictMath para el resto de las funciones matemáticas como continuac ión de HIVE-23133
Corrección de errores	Se corrigió la incompatibilidad comodín en los HMS get_partitions_by_filter y get_num_p artitions_by_filter causada por la portabilidad de HIVE-22900 en EMR Hive 6.4.0 APIs

Tipo	Descripción
Corrección de errores	HIVE-26736: fallo de autorización para las vistas anidadas que tienen la cláusula WITH
Corrección de errores	HIVE-22416: faltan registros de operacion es relacionadas con MR cuando la ejecución paralela está habilitada
Corrección de errores	HIVE-19653: se ha presionado incorrectamente el predicado de groupby con conjuntos de agrupamiento
Corrección de errores	HIVE-22094: las consultas fallaban con: ClassCastException hive.ql.exec.vector. DecimalColumnVectorno se puede convertir en hive.ql.exec.Vector.Decimal64 ColumnVector
Corrección de errores	HIVE-26340: el operador PTF vectorizado falla si la consulta tiene la función de ventana en mayúsculas
Corrección de errores	HIVE-26184: COLLECT_SET con GROUP BY es muy lento cuando algunas claves están muy sesgadas
Corrección de errores	ClassCastException HIVE-26373: al leer las marcas de tiempo de una tabla con datos de Avro HBase
Corrección de errores	HIVE-26388: ClassCastException cuando no hay una columna de tipo cadena en la tabla fuente de la consulta CTAS Actualice HIVE-26172: Hive: Actualice Ant a la versión 1.10.11 debido a los CVE-2021-36373 y CVE-2021-36374

Tipo	Descripción
Corrección de errores	HIVE-26114: la corrección de la conexión jdbc hiveserver2 mediante el comando dfs con espacio de prefijo provocará una excepción
Corrección de errores	HIVE-26396: la función trunc tiene un problema con la intercepción de precisión y el resultado tiene muchos 0
Corrección de errores	HIVE-26446: TablesWritten no HiveProto LoggingHook se puede rellenar el campo de las tablas particionadas.
Corrección de errores	HIVE-26639: y no debería basarse en el juego de caracteres predeterminado ConstantV ectorExpression ExplainTask
Corrección de errores	HIVE-22670: ArrayIndexOutOfBoundsExcept ion cuando se utiliza un lector vectorizado para leer un archivo de parquet
Corrección de errores	HIVE-23607: problema de permiso: la creación de una vista en otra vista se lleva a cabo correctamente, pero la modificación de la vista falla
Corrección de errores	HIVE-25498: consulta con más de 31 funciones distintas devuelve un resultado incorrecto
Corrección de errores	HIVE-25780: crea más de 64 conjuntos de agrupamiento II DistinctExpansion
Corrección de errores	HIVE-23868: especificación de la función de creación de ventanas: admite 0 anteriores o siguientes

Tipo	Descripción
Corrección de errores	HIVE-24539: la generación del esquema debe respetar el delimitador de columnas OrcInputFormat
Corrección de errores	HIVE-23476: LLAP: preasigne también las arenas para el caso del mapa
Corrección de errores	HIVE-25806: Posible filtración - Parquet, LLAP IO LlapCacheAwareFs
Corrección de errores	HIVE-23498: Desactiva el método de rastreo HTTP en ThriftHttpCliService
Corrección de errores	HIVE-25729: ThriftUnionObjectInspector debe notificarse cuando esté completamente invitado
Corrección de errores	HIVE-23846: se evitan la serialización y deserialización innecesarias de vectores de bits
Corrección de errores	HIVE-24233: excepto la subconsulta arroja un puntero nulo con el cbo desactivado
Corrección de errores	HIVE-24276: Vulnerabilidad de HiveServer 2 loggerconf jsp Cross-Site Scripting (XSS)
Corrección de errores	HIVE-25721: el resultado de la unión externa es incorrecto
Corrección de errores	HIVE-25223: si se selecciona con límite, no se devuelven filas en una tabla no nativa
Corrección de errores	HIVE-25794:: registrar las sentencias en un bucle genera presión en la memoria CombineHiveRecordReader

Tipo	Descripción
Corrección de errores	HIVE-23602: uso del paquete concurrente de Java para el conjunto de identificadores de operación
Corrección de errores	HIVE-24045: no hay registro relacionado con el momento en que se crea la base de datos predeterminada
Corrección de errores	HIVE-24305: el esquema avro decimal no completa correctamente la escala/precisión si el valor está entre comillas
Corrección de errores	HIVE-25040: Drop Database Cascade no puede eliminar las funciones persistentes
Corrección de errores	HIVE-23501: AOOB se produce VectorDes erializeRow cuando los tipos complejos se convierten en tipos primitivos
Corrección de errores	HIVE-23704: el servidor HTTP de Thrift no gestiona correctamente el identificador de autenticación
Corrección de errores	HIVE-23529: el CTAS para uniontype no funciona cuando row_deserialize
Corrección de errores	HIVE-24144: La cadena de entrada devuelve un valor incorrecto getIdentifierQuote HiveDatabaseMetaData
Corrección de errores	HIVE-23850: se permite el PPD cuando el asunto no es una columna con conjuntos de agrupamientos
Corrección de errores	HIVE-25919: ClassCastException al introducir el predicado de una columna booleana HBase StorageHandler

Tipo	Descripción
Corrección de errores	HIVE-25261: Al volver HMSHandler a intentarl o, debería incluir una breve descripción del objetivo MetaException
Corrección de errores	HIVE-24792: posible fuga de subprocesos durante la operación
Corrección de errores	HIVE-23409: Si no se puede volver a abrir TezSession la aplicación o el servicio Timeline está inactivo, se cierra de forma predeterm inada tras volver a intentarlo TezSession SessionPool
Corrección de errores	HIVE-23615: no haga referencia a los punteros nulos en la clase de comandos de Beeline
Corrección de errores	HIVE-24849: se crea un tiempo de espera para un socket de tabla externo cuando la ubicación tiene una gran cantidad de archivos (afecta a la versión 3.1.2)
Corrección de errores	HIVE-25209: consulta SELECT con la función SUM que produce un resultado inesperado
Corrección de errores	HIVE-23666: Se omite checkHashMode la eficiencia cuando un operador de groupby no tiene un conjunto de agrupamientos
Corrección de errores	HIVE-23873: Se produce un error al consultar la tabla Hive Handler con el NPE cuando el CBO está desactivado JDBCStorage
Corrección de errores	HiveStreamingConnection HIVE-24149: no cierra la conexión HMS

Tipo	Descripción
Corrección de errores	HIVE-25561: la tarea finalizada no debería confirmar el archivo (afecta a las versiones 2.x y 3.x).
Corrección de errores	HIVE-25683: Cierre el lector. AcidUtils isRawFormatExpediente
Corrección de errores	HIVE-24294: TezSessionPool las sesiones pueden arrojar AssertionError
Corrección de errores	HIVE-24182: problema de autorización de Ranger con carácter permanente UDFs
Corrección de errores	HIVE-22805: la vectorización con matriz o mapa condicional no está implementada y devuelve un error
Corrección de errores	HIVE-22828: Decimal64: las instrucciones NVL y CASE convierten implícitamente decimal64 en 128
Corrección de errores	HIVE-21398: las columnas que tienen estadísti cas estimadas no deben considerarse claves únicas
Corrección de errores	HIVE-22490: si se agregan jars con caracteres especiales en su ruta, se produce un error
Corrección de errores	HIVE-22700: las compactaciones pueden provocar pérdidas de memoria si no están autorizadas
Corrección de errores	HIVE-22053: el nombre de la función no está normalizado al crear la función

Tipo	Descripción
Corrección de errores	HIVE-22595: las inserciones de particion es dinámicas fallan en la tabla Avro con un esquema externo
Corrección de errores	HIVE-21795: es posible que falte la fila de resumen acumulativo cuando se produce mapjoin en una tabla particionada
Corrección de errores	HIVE-22987: cuando es nulo ClassCast Exception VectorCoalesce DataTypeP hysicalVariation
Corrección de errores	HIVE-22219: la desactivación de un administr ador de nodos impide el reinicio del servicio LLAP
Corrección de errores	HIVE-21793: la CBO recupera las estadísticas de las columnas incluso si hive.stats.fetch.c olumn.stats tiene el valor false
Corrección de errores	HIVE-22163: CBO: al habilitar la CBO, se activa la estimación de estadísticas, incluso cuando la estimación está desactivada
Corrección de errores	HIVE-18735: al crear una tabla se pierde el atributo transaccional
Corrección de errores	HIVE-22433: controlador de almacenam iento JDBC de Hive: se obtuvieron resultados incorrectos de BOOLEAN y TIMESTAMP de la fuente de datos JDBC DataType
Corrección de errores	ObjectStoreHIVE-19430:. cleanNotificationE vents OutOfMemory sobre un gran número de eventos pendientes

Tipo	Descripción
Corrección de errores	HIVE-20785: Nombre de clave incorrecto en el JDBC. DatabaseMetaData getPrimaryKeys método
Corrección de errores	HIVE-16116: Beeline devuelve un NPE cuando beeline.hiveconfvariables={} está en beeline.properties
Corrección de errores	HIVE-20066: hive.load.data.owner se compara con la entidad principal completa
Corrección de errores	HIVE-20489: explicación del plan de bloqueos de consultas
Corrección de errores	HIVE-21033: Si se olvida de cerrar la operación, se cortan 2 salidas adicionales HiveServer
Corrección de errores	HIVE-19888: Advertencia engañosa de «METASTORE_FILTER_HOOK será ignorado» procedente de SessionState
Corrección de errores	HIVE-20303: INSERTE LA PARTICIÓN db.table DE LA TABLA DE SOBRESCRITURA (...) SI NO EXISTE, LANZA InvalidTableException
Corrección de errores	HIVE-16144: CompactionInfo no tiene Equals/hashCode pero se usa en Set
Corrección de errores	HIVE-20818: las vistas creadas con una subconsulta WHERE considerarán las vistas a las que se haga referencia en la subconsulta como una entrada directa
Corrección de errores	HIVE-21005: LLAP: Se filtran más franjas por división ZlibCodecs

Tipo	Descripción
Corrección de errores	HIVE-20771: falla en estructuras vacías. LazyBinarySerDe
Corrección de errores	HIVE-18852: mensaje de error confuso al modificar la validación de la tabla
Corrección de errores	HIVE-21124: HPL/SQL no admite la instrucción CREATE TABLE LIKE
Corrección de errores	HIVE-20935: Se produce un error al cargar el archivo tar del paquete LLAP y provocar un fallo en el inicio del servicio LLAP EC2
Corrección de errores	HIVE-20409: Hive ACID: no limpia el directorio provisional de hdfs Update/delete/merge
Corrección de errores	HIVE-20570: la combinación ALL con hive.optimize.union.remove=true tiene un plan incorrecto
Corrección de errores	HIVE-20421: la entidad de caracteres “\b” no es válida en hive-default.xml.template
Corrección de errores	HIVE-19133: Las métricas de rendimiento de HS2 WebUI por fases no se muestran correctamente
Corrección de errores	HIVE-18977: al enumerar las particiones se obtienen resultados diferentes con JDO y SQL directo
Corrección de errores	HIVE-20034: Anule los cambios en la gestión de excepciones para garantizar la compatibilidad con versiones anteriores MetaStore

Tipo	Descripción
Corrección de errores	HIVE-20672: Al iniciar sesión, el subproceso debería indicar todos los intervalos fijos LlapTaskSchedulerService
Corrección de errores	HIVE-12812: habilita mapred.input.dir.recursive de forma predeterminada para admitir la unión con la función agregada
Corrección de errores	HIVE-20147: la ingesta de flujos de Hive se lleva a cabo mediante el registro sincronizado
Corrección de errores	HIVE-19203: Problema de seguridad de los hilos en HiveMetaStore
Corrección de errores	HIVE-20091: Prueba: Añadir credenciales de seguridad para la salida FileSinkOperator
Corrección de errores	HIVE-16906: Hive ATSHook debe comprobar la existencia de yarn.timeline-service.enabled antes de conectarse a ATS
Corrección de errores	HIVE-20714: SHOW tblproperties para una sola propiedad devuelve el valor de la columna de nombres
Corrección de errores	HIVE-24730: las clases Shims anulan los valores de hive-site.xml y tez-site.xml de forma silenciosa
Corrección de errores	HIVE-22055: el recuento de selecciones devuelve un resultado incorrecto después de cargar los datos del archivo de texto

Amazon EMR 6.11.0: notas de la versión de Hive

Amazon EMR 6.11.0: cambios en Hive

Tipo	Descripción
Mejoras	Se ha agregado compatibilidad con la eliminación de particiones mediante subprocesos múltiples para mejorar el rendimiento de la eliminación de particiones
Mejoras	Se permite leer archivos de consulta de Hive codificados
Mejoras	Tez Shuffle Handler habilitado de forma predeterminada para los trabajos de Hive en Tez
Error	Se agregó una opción para permitir la distribución determinista de las claves a los reductores para corregir un resultado incorrecto cuando hive.groupby.skewindata está habilitado (se informó en HIVE-20220)
Error	Se corrigió un error en el cálculo de las estadísticas cuando se configuraba el nombre de partición predeterminado
Error	Respeta cualquier parámetro de clasificación SSL personalizado que se apruebe cuando el SSL esté configurado de forma inmediata para 2 en un clúster con el cifrado en tránsito activado HiveServer
Portabilidad con versiones anteriores	HIVE-23617: Se corrigieron los problemas de la API de almacenamiento FindBug
Portabilidad con versiones anteriores	HIVE-26408: Vectorización: corrige la desasignación de columnas temporales y evita

Tipo	Descripción
	reutilizar un elemento secundario como salida ConstantVectorExpression
Portabilidad con versiones anteriores	HiveConfig HIVE-23614: Pase siempre a removeTempOr DuplicateFiles
Portabilidad con versiones anteriores	HIVE-23354: Elimine la comprobación de cordura del tamaño del archivo de compareTe mpOr DuplicateFiles
Portabilidad con versiones anteriores	HIVE-20344: Se corrigió el problema de los lanzamientos de la SBA. PrivilegeSynchronizer AccessControlException También se introdujo la propiedad hive.privilege.synchronizer para deshabilitar el sincronizador de privilegios
Portabilidad con versiones anteriores	HIVE-15826: Support configure 'serializ ation.encoding' para todos SerDes
Portabilidad con versiones anteriores	HIVE-18284: corrección de NPE al insertar datos con la cláusula “distribute by” con la optimización de clasificación de dynpart
Portabilidad con versiones anteriores	HIVE-24930: el cortocircuito de Operator. setDone() debido a una operación secundaria no se utiliza en la ruta de código vectorizada (si childSize == 1)
Portabilidad con versiones anteriores	HIVE-24523: La ruta de lectura vectorizada no respeta el SERDEPROPERTIES de la marca de tiempo LazySimpleSerde
Portabilidad con versiones anteriores	HIVE-23265: los conjuntos de filas duplicados se devuelven con el límite y el desplazamiento establecidos

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21492: no se puede leer el archivo de parquet generado con una herramienta de ahorro o personalizada VectorizedParquetR ecordReader
Portabilidad con versiones anteriores	HIVE-22540: Vectorización: no funcionan con columnas de 64 decimales. VectorizedBatchUtil makeLikeColumnVector ()
Portabilidad con versiones anteriores	HIVE-22588: vaciado de las filas restantes para el resto de los conjuntos de agrupamiento al cambiar el modo groupby vectorial
Portabilidad con versiones anteriores	HIVE-22551: BytesColumnVector InitBuffer debe limpiar el vector y la longitud de forma coherente
Portabilidad con versiones anteriores	HIVE-22448: CBO: ampliación del recuento múltiple distinto con una clave group-by
Portabilidad con versiones anteriores	HIVE-22248: solución de problemas persisten tes en las estadísticas
Portabilidad con versiones anteriores	HIVE-22210: la vectorización puede reutilizar las columnas de salida del cálculo involucradas en el filtrado
Portabilidad con versiones anteriores	HIVE-21531: vectorización: todos los hashcodes NULL no se calculan con Murmur3
Portabilidad con versiones anteriores	HIVE-20419: Vectorización: evita la mutación después de haber sido utilizada en una clave de mapa hash VectorPartitionDesc
Portabilidad con versiones anteriores	ClassCastException HIVE-19388 VectorMap JoinCommonOperator : durante la inicialización

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21584: Preparación para Java 11: el cargador de clases de sistema no es Loader URLClass
Portabilidad con versiones anteriores	HIVE-25107: el registro de Classpath debe estar en el nivel DEBUG (#2271)
Portabilidad con versiones anteriores	HIVE-22097: java.util no compatible. ArrayList para java 1.1
Portabilidad con versiones anteriores	HIVE-23938: LLAP: JDK11 - algunos argumentos jvm relacionados con la rotación del archivo de registro de GC ya no se pueden usar
Portabilidad con versiones anteriores	HIVE-26226: exclusión de jdk.tools dep de hive-metastore en upgrade-acid
Portabilidad con versiones anteriores	HIVE-17879: actualización del complemento Datanucleus Maven
Portabilidad con versiones anteriores	HIVE-27004: DateTimeFormatterBuilder # appendZoneText no puede analizar 'UTC+' en versiones de Java superiores a 8
Portabilidad con versiones anteriores	HIVE-16812: no filtra los eventos de eliminación VectorizedOrcAcidRowBatchReader
Portabilidad con versiones anteriores	VectorizedOrcAcidRowBatchReaderHIVE-17917:. computeOffsetAndOptimización de cubos
Portabilidad con versiones anteriores	HIVE-19985: ACID: omisión de la decodificación de las secciones ROW__ID para consultas de solo lectura

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20635: VectorizedOrcAcidRowBatchReader no filtra los eventos de eliminación de los archivos originales
Upgrade	Actualización de Javadoc a 3.3.1
Upgrade	Actualización de Javassist a 3.24.1-GA
Upgrade	Actualización a la versión 2.0.0-M14 apache-directory-server

Nuevas configuraciones

Nombre	Clasificación	Descripción
hive.metastore.fs.drop.partition.threads	hive-site	Número de subprocesos principales en el grupo de subprocesos de la partición desplegable.
hive.metastore.fs.drop.partition.keepalive.time	hive-site	Tiempo en segundos durante el que un hilo asíncrono de partición inactivo (del grupo de subprocesos) esperará a que llegue una nueva tarea antes de finalizar.
hive.metastore.fs.drop.partition.threadpool.max.queue.size	hive-site	Tamaño máximo de cola que se utilizará en el grupo de subprocesos para eliminar particiones del sistema de archivos.
hive.groupby.enable.deterministic.distribution	hive-site	Habilite la distribución determinista de las claves a los reductores. Transmitirá un

Nombre	Clasificación	Descripción
		valor inicial constante al llamar a la función rand utilizada para la partición aleatoria.
hive.privilege.synchronizer	hive-site	Si se deben sincronizar los privilegios de un autorizador externo periódicamente en 2. HiveServer
hive.cli.query.file.encoding	hive-site	Codificación de archivos para todos los tipos de archivos de consulta (archivo de consulta, archivo de consulta de inicio, archivo rc, etc.) proporcionada en los argumentos de la CLI.
hive.emr.tez.shuffle.enabled	hive-site	Los trabajos de Hive en Tez ahora usan tez_shuffle de forma predeterminada en lugar de mapreduce_shuffle como el controlador de reproducción aleatorio predeterminado.

Configuraciones obsoletas

Las siguientes propiedades de la configuración están en desuso por HIVE-23354 y ya no son compatibles con las versiones 6.11.0 y posteriores de Amazon EMR.

Nombre	Valor predeterminado
`hive.mapred.reduce.tasks.speculative.execution`	`false`
`tez.am.speculation.enabled`	`false`

Amazon EMR 6.10.0: notas de la versión de Hive

Amazon EMR 6.10.0: cambios en Hive

Tipo	Descripción
Característica	Habilite los controles de acceso AWS Lake Formation basados en las consultas de Apache Hive (escritura) mediante IAM Passthrough (API HiveCLI/Steps).
Mejora	Deshabilitar la configuración de `hive.log.explain.output` de forma predeterminada para reducir el tamaño del registro
Portabilidad con versiones anteriores	HIVE-26408: Vectorización: corrige la desasignación de columnas temporales y no reutiliza un elemento secundario como resultado ConstantVectorExpression
Portabilidad con versiones anteriores	HIVE-22269: corrección de un error en el recuento de reductores en las consultas de inserción con particiones dinámicas debido a la falta de estadísticas provocada por HIVE-20703.
Portabilidad con versiones anteriores	HIVE-22891: Omite la extracción en el modo de ejecución sin LLAP PartitionDesc CombineHiveRecord
Portabilidad con versiones anteriores	HIVE-23804: adición de una base de datos predeterminada para las tablas específicas de estadísticas de columnas en el esquema de metaalmacén de Hive para que sean compatibles con versiones anteriores

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-25277: Eliminación lenta de particiones de Hive para almacenes de objetos en la nube con un coste caro ListFiles
Portabilidad con versiones anteriores	HIVE-19202: La CBO falló debido a un error. NullPointerException HiveAggregate isBucketedInput()
Portabilidad con versiones anteriores	HIVE-19048: se ignoran los errores de corrección de Beeline Initscript
Portabilidad con versiones anteriores	HIVE-21085: el registro de vistas materializadas inicia una sesión de Tez no externa
Portabilidad con versiones anteriores	HIVE-21675: CREATE VIEW IF NOT EXISTS devuelve un error en lugar de “ACEPTAR” si la vista ya existe. Se trata de una regresión de Hive 2.
Portabilidad con versiones anteriores	HIVE-21646: Prueba: Evita escapar del contexto del registro TezTasks de subprocesos
Portabilidad con versiones anteriores	HIVE-22054: evita las listas recursivas para comprobar si un directorio está vacío
Portabilidad con versiones anteriores	HIVE-16587: NPE al insertar tipos complejos con valores nulos anidados
Portabilidad con versiones anteriores	HIVE-22647: habilita el grupo de sesiones de forma predeterminada
Portabilidad con versiones anteriores	HIVE-13288: mensaje de excepción confuso en .LocalizeResource DagUtils
Portabilidad con versiones anteriores	HIVE-23870: Optimice múltiples conversiones de texto en. WritableHiveCharObjectInspector getPrimitiveJavaObjeto/ HiveCharWritable

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21498: actualización de Thrift a la versión 0.13.0
Portabilidad con versiones anteriores	HIVE-24378: los espacios iniciales y finales no se eliminan antes de la conversión decimal
Portabilidad con versiones anteriores	HIVE-21341: valores predeterminados razonables: hive.server2.idle.operation .timeout y hive.server2.idle.session.timeout son demasiado altos
Portabilidad con versiones anteriores	HIVE-22465: Agregue la configuración ssl TezConfigurationFactory
Portabilidad con versiones anteriores	HIVE-24710: optimización de la iteración del PTF para count(*) para reducir el costo de CPU y E/S
Portabilidad con versiones anteriores	HIVE-15406: consideración de la posibilidad de vectorizar la nueva función “trunc”
Portabilidad con versiones anteriores	HIVE-21541: corrección de los encabezados asf que faltan en HIVE-15406
Portabilidad con versiones anteriores	HIVE-24808: fechas analizadas en caché
Portabilidad con versiones anteriores	HIVE-24746: PTF: se puede optimizar durante el cálculo del alcance TimestampValueBoun daryScanner
Portabilidad con versiones anteriores	HIVE-25059: el evento Alter se convierte para cambiarle el nombre durante la replicación
Portabilidad con versiones anteriores	HIVE-25142: la repetición de códigos en la asignación para unirse a una tabla de hash rápida provoca daños en las claves de gran tamaño

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-23756: se agregaron más restricciones al archivo package.jdo
Portabilidad con versiones anteriores	HIVE-25150: los caracteres de tabulación no se eliminan antes de la conversión decimal, al igual que el carácter de espacio, que se fija como parte de HIVE-24378
Portabilidad con versiones anteriores	HIVE-25093: la UDF de date_format() devuelve la salida solo en la zona horaria UTC
Portabilidad con versiones anteriores	HIVE-25268: date_format udf devuelve resultados incorrectos para fechas anteriores a 1900 si la zona horaria local no es UTC
Portabilidad con versiones anteriores	HIVE-25338: AIOBE en la configuración UDF si la entrada está vacía
Portabilidad con versiones anteriores	HIVE-22400: el minuto UDF con el tiempo devuelve NULL
Portabilidad con versiones anteriores	HIVE-25058: PTF: se puede optimizar durante el cálculo del alcance pt2 - Timestamp ValueBoundaryScanner isDistanceGreater
Portabilidad con versiones anteriores	HIVE-25449: datediff() produce un resultado incorrecto cuando se ejecuta en una tarea de Tez con una zona horaria que no es UTC
Portabilidad con versiones anteriores	HIVE-23688: Vectorización: para una columna de tipo mapa que incluye un valor nulo IndexArrayOutOfBoundsException
Portabilidad con versiones anteriores	HIVE-22247: Hive se lanza cuando la salida de la tarea de la partición está vacía HFile OutputFormat FileNotFoundException

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-25570: Hive debe enviar la ruta URL completa para autorizar el comando insert overwrite location
Portabilidad con versiones anteriores	HIVE-22903: la función row_number() vectoriza da restablece el número de fila después de un lote en caso de una expresión constante en la cláusula de partición
Portabilidad con versiones anteriores	HIVE-25549: resultados incorrectos para la función de ventana con una expresión en la cláusula PARTITION BY u ORDER BY
Portabilidad con versiones anteriores	HIVE-25579: LOAD sobrescribe los anexos en lugar de sobrescribirlos
Portabilidad con versiones anteriores	HIVE-25659: las consultas SQL directas de metaalmacén con IN/(NOT IN) deben dividirse en función del máximo de parámetros permitido s por la base de datos SQL
Portabilidad con versiones anteriores	HIVE-20502: corrección de NPE al ejecutar skewjoin_mapjoin10.q cuando se utilizaban las estadísticas de columnas.
Portabilidad con versiones anteriores	HIVE-25765: la propiedad skip.header.line.c ount omite las filas de cada bloque cuando el tamaño del archivo es mayor FetchOperator
Error	Corrección de NPE al insertarlo en ciertos escenarios cuando `hive.stats.column. autogather` y `hive.groupby.skewi ndata` están habilitados
Error	Corrección de NPE cuando el valor `mapred.tasktracker.expiry.i nterval` no está establecido

Amazon EMR 6.9.0: notas de la versión de Hive

Amazon EMR 6.9.0: cambios en Hive

Tipo	Descripción
Upgrade	Actualización de Jetty a 9.4.48.v20220622
Upgrade	Compatibilidad con Hadoop 3.3.3
Característica	Integración de Amazon EMR Hive con Lake Formation para cargas de trabajo interactivas mediante la API de GCSC.
Característica	Integración de Amazon EMR Hive con Iceberg.
Mejora	Active SSL en HiveServer 2 cuando el cifrado en tránsito esté habilitado mediante las configuraciones de seguridad de Amazon EMR.
Mejora	Habilite el confirmador optimizado de Hive EMRFS de Amazon S3 de forma predeterm inada. Para obtener más información, consulte, Habilitación del confirmador optimizado Hive EMRFS S3.
Mejora	Agregue `HiveHBaseTableInpu tFormatV2` que hereda solo la versión mapeada de para corregir el SPARK-342 10 InputFormat . Configure `hive.hbas e.inputformat.v2` en `true` para usarlo.
Mejoras	Espere a que TezAM se inicie en segundo plano con hive.cli.tez.session.async en lugar de cerrarlo y lanzar uno nuevo inmediatamente. Use `hive.emr.cli.tez.session.op en.timeout` para configurar este tiempo de espera en segundos.

Tipo	Descripción
Mejoras	Agregue la opción hive.conf.restricted.list.append para anexar configuraciones separadas por comas a la lista `hive.conf.restricted.list` de configuraciones restringidas existente.
Mejoras	Se muestra un mensaje de error más claro cuando se produce un error en la consulta de Hive porque la ubicación de la base de datos no está definida.
Portabilidad con versiones anteriores	HIVE-24484: actualización de Hadoop a 3.3.1 y Tez a 0.10.2
Portabilidad con versiones anteriores	HIVE-22398: Elimine la gestión de colas de YARN mediante. ShimLoader
Portabilidad con versiones anteriores	HIVE-23190: LLAP: modificar para pasar un objeto del sistema de archivos a. IndexCache TezSpillRecord
Portabilidad con versiones anteriores	HIVE-22185: El HADOOP-15832 provocará problemas con las pruebas que utilicen clústeres. MiniYarn
Portabilidad con versiones anteriores	HIVE-21670: sustitución de `mockito-all` por la dependencia de `mockito-core` .
Portabilidad con versiones anteriores	HIVE-24542: preparación de Guava para las actualizaciones.
Portabilidad con versiones anteriores	HIVE-23751:: Reemplace el método para alinearlo después del HADOOP-16582. QTest `#mkdirs()` `ProxyFileSystem`
Portabilidad con versiones anteriores	HIVE-21603: preparación de Java 11: actualización de la versión powermock.

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-24083: error de `hcatalog` en Hadoop 3.3.0: se necesita un tipo de autenticación.
Portabilidad con versiones anteriores	HIVE-24282: muestra que las columnas no deben ordenar las columnas de salida a menos que se mencione explícitamente.
Portabilidad con versiones anteriores	HIVE-20656: valores predeterminados razonables: las configuraciones de memoria de agregación de asignaciones son demasiado agresivas.
Portabilidad con versiones anteriores	HIVE-25443: Arrow SerDe no puede serializa r/deserializar tipos de datos complejos cuando hay más de 1024 valores
Portabilidad con versiones anteriores	HIVE-19792: actualización de ORC a 1.5.2 y habilitación de pruebas de evolución del esquema decimal_64.
Portabilidad con versiones anteriores	HIVE-20437: gestión de la evolución del esquema a partir de valores flotantes, dobles y decimales.
Portabilidad con versiones anteriores	HIVE-21987: Hive no puede leer el int32 de Parquet anotado con decimales.
Portabilidad con versiones anteriores	HIVE-20038: las consultas de actualización en tablas no agrupadas y particionadas generan un NPE.

Amazon EMR 6.9.0: problemas conocidos de Hive

- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación

dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.

Amazon EMR 6.8.0: notas de la versión de Hive

Amazon EMR 6.8.0: cambios en Hive

Tipo	Descripción
Mejoras	Reducción de las llamadas al sistema de archivos en el comando msck. Mejoras en el rendimiento (entre 15 y 20 veces en más de 10 000 particiones)
Portabilidad con versiones anteriores	HIVE-20678: Hive debe implementarse para garantizar la compatibilidad HBase TableOutputFormat HiveOutputFormat
Portabilidad con versiones anteriores	HIVE-21040: msck incluye archivos innecesarios en el último nivel del árbol de directorios
Portabilidad con versiones anteriores	HIVE-21460: al cargar los datos seguidos de una consulta de selección *, se obtienen resultados incorrectos
Portabilidad con versiones anteriores	HIVE-21660: se obtiene un resultado incorrecto cuando se utiliza la vista de unión total y posterior con explode
Portabilidad con versiones anteriores	HIVE-22505: se debe a una selección incorrecta de un operador vectorizado ClassCast Exception

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-22513: la propagación constante de la columna proyectada en las operaciones de filtrado puede provocar resultados incorrectos
Portabilidad con versiones anteriores	HIVE-23435: faltan filas en el resultado de la unión exterior completa
Portabilidad con versiones anteriores	HIVE-24209: conversión incorrecta de los argumentos de búsqueda para la operación NOT BETWEEN cuando la vectorización está habilitada
Portabilidad con versiones anteriores	HIVE-24934: la anotación no es necesaria en Generic Check VectorizedExpressions UDFSQCount
Portabilidad con versiones anteriores	HIVE-25278: HiveProjectJoinTransposeRule puede realizar transformaciones no válidas con expresiones en ventanas
Portabilidad con versiones anteriores	HIVE-25505: resultados incorrectos con el encabezado. skip.header.line.count si la primera línea está en blanco
Portabilidad con versiones anteriores	HIVE-26080: actualización de accumulo-core a 1.10.1
Portabilidad con versiones anteriores	HIVE-26235: condición OR de la columna binaria devuelve un resultado vacío
Error	Corrige varios registros de advertencia de enlaces SLF4 J en stderr durante el lanzamien to

Tipo	Descripción
Error	Se ha corregido un error en la consulta SHOW TABLE EXTENDED que provocaba un error de FS incorrecto cuando la partición y la tabla se encontraban en sistemas de archivos diferente s.

Amazon EMR 6.8.0: problemas conocidos de Hive

- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.

Amazon EMR 6.7.0: notas de la versión de Hive

Amazon EMR 6.7.0: cambios en Hive

Tipo	Descripción
Característica	Integración de Amazon EMR Hive con. LakeFormation
Característica	Registro de auditoría adicional para el confirmador Hive EMRFS optimizado para Amazon S3. Configuración de Hive: `hive.blobstore.output-commi tter.logging` , valor predeterminado: `false`
Característica	El directorio de destino eliminado al insertarlo se sobrescribe con el resultado de selección

Tipo	Descripción
	vacío en una tabla/partición estática sin particiones para que se comporte de manera similar a Hive 2.x. Configuración de Hive: `hive.emr.iow.clean.target.dir` , valor predeterminado: `false`
Error	Se corrigió un error de consulta intermite nte al utilizar el confirmador de Hive EMRFS optimizado para Amazon S3 con clasificación de buckets de particiones.
Upgrade	Versión de Hive actualizada a la versión 3.1.3. Consulte las notas de la versión 3.1.3 de Apache Hive para obtener más información.
Upgrade	Se actualizó Parquet a la versión 1.12.2.
Portabilidad con versiones anteriores	HIVE-20065: el metaalmacén no debería confiar en Jackson 1.x
Portabilidad con versiones anteriores	HIVE-20071: se migra a Jackson 2.x y se evita su uso
Portabilidad con versiones anteriores	HIVE-20607: se TxnHandler debe utilizar para PreparedStatement ejecutar consultas SQL directas
Portabilidad con versiones anteriores	HIVE-20740: Elimina el bloqueo global en el método .setConf ObjectStore
Portabilidad con versiones anteriores	HIVE-20961: retirar la implementación de NVL
Portabilidad con versiones anteriores	HIVE-22059: el jar hive-exec no contiene la biblioteca jackson (fasterxml)
Portabilidad con versiones anteriores	HIVE-22351: Se corrige el uso incorrecto de subprocesos en ObjectStore TestObjectStore

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-23534: NPE en #invoke cuando se captura sin mensaje RetryingMetaStoreClient MetaException
Portabilidad con versiones anteriores	HIVE-24048: armonización de los componentes de Jackson a la última versión 2.10 de Hive
Portabilidad con versiones anteriores	HIVE-24768: uso de jackson-bom en todas partes para reemplazar la versión
Portabilidad con versiones anteriores	HIVE-24816: actualización de Jackson a la versión 2.10.5.1 o 2.11.0+ gracias a la versión CVE-2020-25649
Portabilidad con versiones anteriores	HIVE-25971: el cierre de la tarea de Tez se retrasa debido a que el grupo de subprocesos en caché no está cerrado
Portabilidad con versiones anteriores	HIVE-26036: NPE causada por get () in MTable ObjectStore

Amazon EMR 6.7.0: problemas conocidos de Hive

- Las consultas con funciones de ventana en la misma columna que la de unión pueden provocar transformaciones no válidas, como se indica en HIVE-25278, y provocar resultados incorrectos o errores en las consultas. Una solución alternativa sería deshabilitar la CBO en el nivel de consulta para este tipo de consultas. La corrección estará disponible en una versión de Amazon EMR posterior a la 6.7.0. Para obtener más información, ponte en contacto con el servicio de asistencia. AWS
- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva

característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.

Amazon EMR 6.6.0: notas de la versión de Hive

Amazon EMR 6.6.0: cambios en Hive

Tipo	Descripción
Upgrade	Actualización de Parquet a la versión 1.12.1.
Upgrade	Actualización de la versión de jars jetty a 9.4.43.v20210629
Error	Se ha corregido un problema que provocaba que Hive se instalara en todos los nodos principales o de tareas cuando la LLAP estaba habilitada en un clúster de Hive.
Portabilidad con versiones anteriores	HIVE-25942: actualización de commons-io a la versión 2.8.0 debido a CVE-2021-29425
Portabilidad con versiones anteriores	HIVE-25726: actualización de velocidad a 2.3 debido a CVE-2020-13936
Portabilidad con versiones anteriores	HIVE-25680: autoriza a la API HiveMetastore del servidor #get_table_meta a utilizar cualquier a de los HiveMetastore modelos de autorizac ión.
Portabilidad con versiones anteriores	HIVE-25554: actualización de la versión Arrow a la 0.15
Portabilidad con versiones anteriores	HIVE-25242: la consulta funciona muy lentamente con vectorized.adaptor = chosen
Portabilidad con versiones anteriores	HIVE-25085: Los MetaStore clientes ya no se comparten entre sesiones.

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-24827: la consulta de agregación de Hive devuelve resultados incorrectos para los archivos que no son de texto.
Portabilidad con versiones anteriores	HIVE-24683: Las cuñas Hadoop23 son propensas a sufrir NPE en rutas getFileId inexistentes
Portabilidad con versiones anteriores	HIVE-24656: CBO falla cuando las consultas son nulas en los tipos de asignaciones y matrices
Portabilidad con versiones anteriores	HIVE-24556: DefaultGraphWalker Optimizado para casos sin nieto
Portabilidad con versiones anteriores	HIVE-24408: actualización de Parquet a 1.11.1
Portabilidad con versiones anteriores	HIVE-24391: Se corrigen los errores de FIX en la rama 3.1 TestOrcFile
Portabilidad con versiones anteriores	HIVE-24362: el procesamiento de árboles AST no es óptimo para árboles con una gran cantidad de nodos
Portabilidad con versiones anteriores	HIVE-24316: actualización de ORC de 1.5.6 a 1.5.8 en branch-3.1
Portabilidad con versiones anteriores	HIVE-24307: Beeline con los parámetros property-file y -e está fallando
Portabilidad con versiones anteriores	HIVE-24245: PTF vectorizado con recuento y partición distinta que produce resultados incorrectos.
Portabilidad con versiones anteriores	HIVE-24224: corrección del error de omitir el encabezado o pie de página de Hive en Tez en un archivo comprimido

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-24157: el modo estricto falla en el sistema numérico CAST timestamp ↔
Portabilidad con versiones anteriores	HIVE-24113: NPE en formato genérico UDFTo UnixTimeStamp
Portabilidad con versiones anteriores	HIVE-23987: actualización de la versión Arrow a 0.11.0
Portabilidad con versiones anteriores	HIVE-23972: adición del ID de cliente externo al cliente externo LLAP
Portabilidad con versiones anteriores	HIVE-23806: evite borrar los estados estadísti cos de las columnas en todas las particiones en caso de que se amplíe el esquema. Esto mejora el tiempo de ejecución de la instrucción alter table add columns.
Portabilidad con versiones anteriores	HIVE-23779: La BasicStatsTask información no se imprime en la consola beeline
Portabilidad con versiones anteriores	HIVE-23306: el comando RESET no funciona si System.getProperty ha establecido una configuración
Portabilidad con versiones anteriores	HIVE-23164: el servidor no está cerrado correctamente debido a subprocesos sin daemon
Portabilidad con versiones anteriores	HIVE-22967: compatibilidad de hive.relo adable.aux.jars.path para Hive en Tez
Portabilidad con versiones anteriores	HIVE-22934: el registro interactivo del servidor Hive registra el flujo de errores
Portabilidad con versiones anteriores	HIVE-22901: la sustitución de variables puede provocar un OOM en las referencias circulares

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-22769: resultados de consulta incorrect os y error de consulta durante la generación dividida de archivos de texto comprimidos
Portabilidad con versiones anteriores	HIVE-22716: Leer a es ilegal ByteBuffer ParquetFooterInputFromCache
Portabilidad con versiones anteriores	HIVE-22648: actualización de Parquet a la versión 1.11.0
Portabilidad con versiones anteriores	HIVE-22640: Decimal64: cuando el tipo de columna de partición es decimal ColumnVector ClassCastException
Portabilidad con versiones anteriores	HIVE-22621: caso de prueba inestable : .testSigning TestLlapSignerImpl
Portabilidad con versiones anteriores	HIVE-22533: se corrigen posibles vulnerabi lidades de la interfaz de usuario web del daemon LLAP
Portabilidad con versiones anteriores	HIVE-22532: Es posible que PTFPPD supere el límite de forma incorrecta a través de la función Rank/ DenseRank
Portabilidad con versiones anteriores	HiveProtoLoggingHook HIVE-22514: puede consumir mucha memoria
Portabilidad con versiones anteriores	HIVE-22476: la función datediff de Hive proporcionó resultados inconsistentes cuando hive.fetch.task.conversion se estableció en none
Portabilidad con versiones anteriores	HIVE-22429: las tablas agrupadas migradas con bucketing_version 1 en Hive 3 utilizan bucketing_version 2 para las inserciones

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-22412: lanza NPE cuando expliques StatsUtils
Portabilidad con versiones anteriores	HIVE-22360: MultiDelimitSerDe devuelve resultados incorrectos en la última columna cuando el archivo cargado tiene más columnas que las del esquema de la tabla
Portabilidad con versiones anteriores	HIVE-22332: Hive debe garantizar una configuración de evolución del esquema válida desde el ORC-540
Portabilidad con versiones anteriores	HIVE-22331: unix_timestamp sin argumento devuelve la marca de tiempo en milisegundos en lugar de segundos
Portabilidad con versiones anteriores	HIVE-22275:. OperationManager queryIdOperationno limpia correctamente varios ID de consulta
Portabilidad con versiones anteriores	HIVE-22273: la comprobación de acceso falla cuando se elimina un directorio temporal
Portabilidad con versiones anteriores	HIVE-22270: actualización de commons-io a 2.6
Portabilidad con versiones anteriores	HIVE-22241: implementación de la UDF para interpretar la fecha y la marca de tiempo con su representación interna y el calendario híbrido gregoriano-juliano
Portabilidad con versiones anteriores	HIVE-22241: implementación de la UDF para interpretar la fecha y la marca de tiempo con su representación interna y el híbrido gregoriano-juliano

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-22232: NPE cuando hive.order.columna lignment se establece en false
Portabilidad con versiones anteriores	HIVE-22231: la consulta de Hive de gran tamaño mediante Knox falla y Broken Pipe Write ha fallado
Portabilidad con versiones anteriores	HIVE-22221: cliente externo de Llap: es necesario reducir el #getSplits LlapBaseI nputFormat
Portabilidad con versiones anteriores	HIVE-22208: el nombre de la columna con la palabra clave reservada no se escapa cuando se vuelve a escribir la consulta que incluye unir una tabla con una columna de máscara
Portabilidad con versiones anteriores	HIVE-22197: Common Merge se une y lanza una excepción de reparto de clases.
Portabilidad con versiones anteriores	HIVE-22170: from_unixtime y unix_timestamp deberían usar la zona horaria de la sesión del usuario
Portabilidad con versiones anteriores	HIVE-22169: Tez: SplitGenerator intenta buscar archivos de planes que no existirían para Tez
Portabilidad con versiones anteriores	HIVE-22168: eliminación de registros muy costosos de la ruta de acceso a la memoria caché
Portabilidad con versiones anteriores	HIVE-22161: UDF: se sincroniza en org.apach e.hadoop.hive.ql.udf. FunctionRegistry UDFType clase

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-22120: Se corrigen los resultados o ArrayOutOfBound excepciones incorrectos en las uniones de la izquierda del mapa exterior en condiciones de límite específicas
Portabilidad con versiones anteriores	HIVE-22115: se impide la creación de un agregador de enrutamiento de consultas si la propiedad está establecida en falso
Portabilidad con versiones anteriores	HIVE-22113: Impide que el LLAP se apague en aplicaciones relacionadas AMReporter RuntimeException
Portabilidad con versiones anteriores	HIVE-22106: se elimina la sincronización entre consultas para la evaluación de particiones
Portabilidad con versiones anteriores	HIVE-22099: Varias fechas relacionadas con UDFs fechas no pueden gestionar correctam ente las fechas julianas desde HIVE-20007
Portabilidad con versiones anteriores	HIVE-22037: debería registrarse al apagarse debido al OOM HS2
Portabilidad con versiones anteriores	HIVE-21976: El desfase debería ser nulo en lugar de cero en Calcita HiveSortLimit
Portabilidad con versiones anteriores	HIVE-21924: divide los archivos de texto incluso si existe un encabezado o pie de página
Portabilidad con versiones anteriores	HIVE-21913: Los UDTFGet splits genéricos deben tratar los nombres de usuario de la misma forma que los LLAP
Portabilidad con versiones anteriores	HIVE-21905: Los genéricos han mejorado en toda su clase FetchOperator

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21902: 2 UI: HiveServer se necesita un encabezado de respuesta X-Frame-Options
Portabilidad con versiones anteriores	HIVE-21888: establece el valor predeterminado de hive.parquet.timestamp.skip.conversion en verdadero
Portabilidad con versiones anteriores	HIVE-21868: vectorizar CAST...FORMAT
Portabilidad con versiones anteriores	HIVE-21864: #closeAll LlapBaseInputFormat
Portabilidad con versiones anteriores	HIVE-21863: mejora de la conversión del tipo vectorizador para la expresión WHEN
Portabilidad con versiones anteriores	HIVE-21862: ORC ppd produce un resultado incorrecto con la marca de tiempo
Portabilidad con versiones anteriores	HIVE-21846: Crea un hilo en TeZam que busque métricas periódicamente LlapDaemon
Portabilidad con versiones anteriores	HIVE-21837: MapJoin arroja una excepción cuando la columna seleccionada tiene valores completamente nulos
Portabilidad con versiones anteriores	HIVE-21834: se evitan llamadas innecesarias para simplificar las condiciones del filtro
Portabilidad con versiones anteriores	HIVE-21832: nuevas métricas para obtener el tiempo promedio queue/serving/response
Portabilidad con versiones anteriores	HIVE-21827: Múltiples llamadas en el método «no se procesan» SemanticAnalyzer getTableObject ByName
Portabilidad con versiones anteriores	HIVE-21822: Exponga las métricas mediante un nuevo método de API LlapDaemon

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21818: CBO: Copiar tiene tráfico de metastore TableRelOptHiveTable
Portabilidad con versiones anteriores	HIVE-21815: las estadísticas del archivo ORC se analizan dos veces
Portabilidad con versiones anteriores	HIVE-21805: 2: Usa el método rápido HiveServer ShutdownHookManager APIs
Portabilidad con versiones anteriores	HIVE-21799: se activa, cuando la clave de unión está NullPointerException en DynamicPartitionPruningOptimization la columna de agregación
Portabilidad con versiones anteriores	HIVE-21794: Agregue parámetros de vista materializada a sqlStdAuth SafeVarNameRegexes
Portabilidad con versiones anteriores	HIVE-21768: JDBC: eliminación del prefijo de unión predeterminado para las consultas UNION no incluidas
Portabilidad con versiones anteriores	HIVE-21746: ArrayIndexOutOfBoundsException durante la unión de hash particionada dinámicamente, con la CBO desactivada
Portabilidad con versiones anteriores	HIVE-21717: no se puede cambiar el nombre del directorio en la tarea de movimiento.
Portabilidad con versiones anteriores	HIVE-21685: simplificación incorrecta en una consulta con varias cláusulas IN
Portabilidad con versiones anteriores	HIVE-21681: la descripción del formateado muestra información incorrecta sobre varias claves principales

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21651: mueva el servidor protobuf a hive-exec.
Portabilidad con versiones anteriores	HIVE-21619: tipo de marca de tiempo de impresión sin precisión en instrucción explain de SQL ampliada
Portabilidad con versiones anteriores	HIVE-21592: no se muestra cuando la expresión contiene CONCAT OptimizedSql
Portabilidad con versiones anteriores	HIVE-21576: introducción de CAST...FORMAT y lista limitada de formatos de fecha y hora SQL:2016
Portabilidad con versiones anteriores	HIVE-21573: el transporte binario ignorará la entidad principal si la autenticación está configurada en delegationToken
Portabilidad con versiones anteriores	HIVE-21550: TestObjectStore las pruebas son dudosas: no se pudo obtener un bloqueo en el tiempo solicitado
Portabilidad con versiones anteriores	HIVE-21544: La propagación constante corrompe las expresiones durante el plegado coalesce/case/when
Portabilidad con versiones anteriores	HIVE-21539: GroupBy + donde una cláusula de la misma columna provoca una reescritura incorrecta de la consulta
Portabilidad con versiones anteriores	HIVE-21538: Beeline: origen de la contraseña, aunque el lector de la consola no pasó al parámetro de conexión
Portabilidad con versiones anteriores	HIVE-21509: el LLAP puede almacenar en caché los vectores de columnas dañados y devolver un resultado de consulta incorrecto

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21499: no debería eliminar la función del registro si el comando create falló con AlreadyExistsException
Portabilidad con versiones anteriores	HIVE-21496: el dimensionamiento automático del búfer sin ordenar puede desbordarse
Portabilidad con versiones anteriores	HIVE-21468: distinción entre mayúsculas y minúsculas en los nombres de los identific adores del controlador de almacenamiento JDBC
Portabilidad con versiones anteriores	HIVE-21467: eliminación de las importaciones obsoletas de junit.framework.Assert
Portabilidad con versiones anteriores	HIVE-21435: LlapBaseInputFormat debe obtener el número de tarea de la configuración TASK_ATTEMPT_ID, si está presente, durante la compilación SubmitWorkRequestProto
Portabilidad con versiones anteriores	HIVE-21389: la distribución de Hive pierde javax.ws.rs-api.jar después de HIVE-21247
Portabilidad con versiones anteriores	HIVE-21385: se permite deshabilitar la transferencia de datos no separables a orígenes de JDBC
Portabilidad con versiones anteriores	HIVE-21383: controlador de almacenamiento JDBC: uso del catálogo y el esquema para recuperar tablas si se especifica
Portabilidad con versiones anteriores	HIVE-21382: optimización de reducción de grupos por claves: las claves no se reducen en query23

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21362: adición de un formato de entrada y propiedad SerDe para leer los archivos protobuf.
Portabilidad con versiones anteriores	HIVE-21340: CBO: Elimine las columnas que no sean clave y que se introduzcan en un SemiJoin
Portabilidad con versiones anteriores	HIVE-21332: purga de los búferes no bloqueados en lugar de los bloqueados
Portabilidad con versiones anteriores	HIVE-21329: tamaño del búfer de salida desordenado en tiempo de ejecución de Tez personalizado según la canalización del operador
Portabilidad con versiones anteriores	HIVE-21295: convertirá la fecha en cadena utilizando la convención de Hive StorageHandler
Portabilidad con versiones anteriores	HIVE-21294: Vectorización: 1 reductor Shuffle puede omitir las funciones hash del objeto
Portabilidad con versiones anteriores	HIVE-21255: Eliminar QueryConditionBuilder JdbcStorageHandler
Portabilidad con versiones anteriores	HIVE-21253: Support en JDBC DB2 StorageHandler
Portabilidad con versiones anteriores	HIVE-21232: LLAP: adición de un proveedor de doble afinidad compatible con cache-miss
Portabilidad con versiones anteriores	HIVE-21214 MoveTask : Utilice TemptID en lugar del tamaño del archivo para la deduplicación de archivos compareTempOr Duplicate Files

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-21184: adición de instrucción explain y plan CBO con formato explain con información sobre los costos
Portabilidad con versiones anteriores	HIVE-21182: se evita configurar Hive Scratch Dir durante la planificación
Portabilidad con versiones anteriores	HIVE-21171: se evita la creación de directorios scratch para Tez si RPC está activado
Portabilidad con versiones anteriores	HIVE-21126: Permite consultas a nivel de sesión en #getSplit LlapBaseInputFormat
Portabilidad con versiones anteriores	HIVE-21107: error “No se puede encontrar el campo” durante la unión de hash particionada dinámicamente
Portabilidad con versiones anteriores	HIVE-21061: La consulta CTAS falla si la fuente está vacía IllegalStateException
Portabilidad con versiones anteriores	HIVE-21041: NPE, al obtener el esquema del plan lógico ParseException
Portabilidad con versiones anteriores	HIVE-21013: no se pudo encontrar la columna de partición en Oracle JdbcStorageHandler
Portabilidad con versiones anteriores	HIVE-21006: Amplíe SharedWorkOptimizer para eliminar las semiuniones cuando haya una oportunidad de reutilización
Portabilidad con versiones anteriores	HIVE-20992: división de la configuración hive.metastore.dbaccess.ssl.properties en configuraciones más significativas
Portabilidad con versiones anteriores	HIVE-20989: JDBC: el registro + puede bloquear el progreso de las consultas durante el modo de suspensión GetOperationStatus

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20988: resultados incorrectos para las consultas agrupadas por grupos con la clave principal en varias columnas
Portabilidad con versiones anteriores	HIVE-20985: si las entradas del operador seleccionado son columnas temporales, la vectorización puede reutilizar algunas de ellas como salida
Portabilidad con versiones anteriores	HIVE-20978: «hive.jdbc.*» debería añadirse a sqlStdAuth SafeVarNameRegexes
Portabilidad con versiones anteriores	HIVE-20953: eliminación de una función del registro de funciones cuando no se puede agregar al metaalmacén cuando se crea.
Portabilidad con versiones anteriores	HIVE-20952: VectorizationContext Limpiar .java
Portabilidad con versiones anteriores	HIVE-20951: LLAP: establecimiento de Xms siempre en un 50 %
Portabilidad con versiones anteriores	HIVE-20949: mejora de la estimación de la cardinalidad del PKFK en la planificación física
Portabilidad con versiones anteriores	HIVE-20944: no se validan las estadísticas durante la compilación de la consulta
Portabilidad con versiones anteriores	HIVE-20940: casos puente en los que la resolución tipográfica de Calcite es más estricta que la de Hive.
Portabilidad con versiones anteriores	HIVE-20937: la consulta de JDBC de Postgre falla: “LIMIT no puede ser negativo”

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20926: la sugerencia de reducción de semiuniones falla cuando las entradas del filtro de floración son altas o cuando no hay estadísticas
Portabilidad con versiones anteriores	HIVE-20920: uso de las restricciones de SQL para mejorar el algoritmo de reordenamiento de uniones
Portabilidad con versiones anteriores	HIVE-20918: indicador para habilitar o deshabilitar la transferencia de datos desde Calcite a una conexión de JDBC
Portabilidad con versiones anteriores	HIVE-20915: habilitación de la disponibilidad de optimización dinámica de las particiones de clasificación disponible para HO y MR
Portabilidad con versiones anteriores	HIVE-20910: la inserción en una tabla agrupada falla debido a la optimización dinámica de la clasificación de particiones
Portabilidad con versiones anteriores	HIVE-20899: el URI de Keytab para el servicio LLAP YARN es restrictivo y solo admite HDFS
Portabilidad con versiones anteriores	HIVE-20898: para las funciones relacionadas con el tiempo, los argumentos no se pueden convertir en un tipo que no acepte valores NULL
Portabilidad con versiones anteriores	HIVE-20881: la propagación constante simplifica en exceso las proyecciones
Portabilidad con versiones anteriores	HIVE-20880: actualización del valor predeterminado de hive.stats.filter.in.min.ratio

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20873: Utilice el hash Murmur para reducir la colisión del hash VectorHas hKeyWrapperTwoLong
Portabilidad con versiones anteriores	HIVE-20868: La unión de SMB falla de forma intermitente cuando hay una operación secundaria TezDummyOperator getFinalOp MapRecordProcessor
Portabilidad con versiones anteriores	HIVE-20853: Exponga .registerDag en la API llap daemon ShuffleHandler
Portabilidad con versiones anteriores	HIVE-20850: si es posible, coloque mayúsculas y minúsculas para pasar de las proyecciones a las tablas de dimensiones
Portabilidad con versiones anteriores	HIVE-20842: se corrige la lógica introducida en HIVE-20660 para estimar las estadísticas de la instrucción group by
Portabilidad con versiones anteriores	HIVE-20839: error “No se puede encontrar el campo” durante la unión de hash particionada dinámicamente
Portabilidad con versiones anteriores	HIVE-20835: la interacción entre las restricci ones y la reescritura del MV puede crear un bucle en Calcite Planner
Portabilidad con versiones anteriores	HIVE-20834: Las entradas de la colmena guardan una referencia a una consulta en caché QueryResultCache SemanticAnalyzer
Portabilidad con versiones anteriores	HIVE-20830: error de aserción de consulta de rango en algunos casos JdbcStorageHandler
Portabilidad con versiones anteriores	HIVE-20829: la división de rango arroja un NPE JdbcStorageHandler

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20827: resultados incoherentes para matrices vacías
Portabilidad con versiones anteriores	HIVE-20826: Mejore la HiveSemiJoin regla para convertir unir + agrupar por el lado izquierdo en semiunir por la izquierda
Portabilidad con versiones anteriores	HIVE-20821: reescribe SUM0 en la combinación SUM + COALESCE
Portabilidad con versiones anteriores	HIVE-20815: .next no comerá ninguna excepción JdbcRecordReader
Portabilidad con versiones anteriores	HIVE-20813: udf to_epoch_milli también debe admitir la marca de tiempo sin zona horaria.
Portabilidad con versiones anteriores	HIVE-20804: mejoras adicionales en la optimización de agrupación con restricciones
Portabilidad con versiones anteriores	HIVE-20792: al insertar una marca de tiempo con zonas, se truncan los datos
Portabilidad con versiones anteriores	HIVE-20788: la reducción extendida de SJ puede hacer retroceder incorrectamente las columnas al crear filtros
Portabilidad con versiones anteriores	HIVE-20778: es posible que el reordenamiento de las uniones no se active si todas las uniones del plan se crean mediante una lógica de descorrelación
Portabilidad con versiones anteriores	HIVE-20772: registro de los contadores de CPU por tarea en LLAP
Portabilidad con versiones anteriores	HIVE-20768: adición de la UDF de ventana de saltos de tamaño constante

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20767: varios proyectos entre operadores de unión pueden afectar al reordenamiento de las uniones mediante restricciones
Portabilidad con versiones anteriores	HIVE-20762: el intervalo de limpieza de NOTIFICATION_LOG está codificado en 60 segundos y es demasiado pequeño
Portabilidad con versiones anteriores	HIVE-20761: la tabla de secuencia de notificaciones seleccionada para actualizarla tiene un intervalo de reintentos y un recuento de reintentos demasiado pequeño
Portabilidad con versiones anteriores	HIVE-20751: actualización de la versión Arrow a la 0.10.0
Portabilidad con versiones anteriores	HIVE-20746: HiveProtoHookLogger no cierra el archivo al final del día.
Portabilidad con versiones anteriores	HIVE-20744: uso de las restricciones de SQL para mejorar el algoritmo de reordenamiento de uniones
Portabilidad con versiones anteriores	HIVE-20740: Elimina el bloqueo global en el método .setConf. ObjectStore Esta selección es compatible con HIVE-20740, diseñado para Hive 3.2 y 4.x, hasta 3.1.x
Portabilidad con versiones anteriores	HIVE-20734: Beeline: cuando beeline-site.xml existe y la CLI de Hive redirige a Beeline, debe usar el nombre de usuario del sistema o la contraseña ficticia en lugar de solicitar uno
Portabilidad con versiones anteriores	HIVE-20731: se debe autorizar la entrada al archivo de almacén de claves JdbcStorageHandler

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20720: adición de la opción de columna de partición al controlador de JDBC
Portabilidad con versiones anteriores	HIVE-20719: la instrucción SELECT falla después de UPDATE con la optimización y vectorización de hive.optimize.sort.dynamic.partition activadas
Portabilidad con versiones anteriores	HIVE-20718: adición del controlador perf cli con restricciones
Portabilidad con versiones anteriores	HIVE-20716: establecimiento del valor predeterminado de hive.cbo.stats.correlated.multi.key.joins en verdadero
Portabilidad con versiones anteriores	HIVE-20712: debería extraer casos profundos HivePointLookupOptimizer
Portabilidad con versiones anteriores	HIVE-20710: el plegado constante no puede crear constantes nulas sin tipos
Portabilidad con versiones anteriores	HIVE-20706: external_jdbc_table2.q falla de forma intermitente
Portabilidad con versiones anteriores	HIVE-20704: Amplíe para admitir otras funciones HivePreFilteringRule
Portabilidad con versiones anteriores	HIVE-20703: optimización dinámica de las particiones de clasificación según una decisión basada en el costo
Portabilidad con versiones anteriores	HIVE-20702: cuenta de los gastos generales derivados de las estimaciones basadas en la estructura de datos durante la selección de mapjoin

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20692: se permite el plegado de expresiones NOT x IS (NOT) [TRUE|FALSE]
Portabilidad con versiones anteriores	HIVE-20691: Corrija org.apache.hadoop.hive.cli. TestMiniLlapCliDriver. testCliDriver[cttl]
Portabilidad con versiones anteriores	HIVE-20682: la ejecución de consultas asíncronas puede fallar si el subproceso cierra sessionHive compartido
Portabilidad con versiones anteriores	HIVE-20676: HiveServer 2: no Privilege Synchronizer está configurado en estado daemon
Portabilidad con versiones anteriores	HIVE-20660: la estimación del grupo por estadísticas podría mejorarse si se limita el número total de filas a la tabla de origen
Portabilidad con versiones anteriores	HIVE-20652: unir dos fuentes de datos diferentes al JdbcStorageHandler controlador jdbc
Portabilidad con versiones anteriores	HIVE-20651: JdbcStorageHandler la contraseña debe estar cifrada
Portabilidad con versiones anteriores	HIVE-20649: gestor de memoria compatible con LLAP para escritores de ORC
Portabilidad con versiones anteriores	HIVE-20648: LLAP: el grupo de vectores por operador debe usar memoria por ejecutor
Portabilidad con versiones anteriores	HIVE-20646: la condición del filtro de particiones no se transfiere a la consulta del metaalmacén si no es nula

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20644: evite exponer información confidencial mediante una excepción de Hive en tiempo de ejecución
Portabilidad con versiones anteriores	HIVE-20636: mejora de la estimación del número de valores nulos después de la unión externa
Portabilidad con versiones anteriores	HIVE-20632: la consulta con la UDF get_splits falla si se crea una vista materializada en la tabla consultada
Portabilidad con versiones anteriores	HIVE-20627: Las consultas asíncronas simultáneas fallan de forma intermitente y provocan una pérdida de memoria LockException
Portabilidad con versiones anteriores	HIVE-20623: trabajo compartido: ampliación del uso compartido de las entradas de la caché de unión de mapas en LLAP
Portabilidad con versiones anteriores	MultiDelimitSerDe HIVE-20619: HiveServer Se incluye en 2 de forma predeterminada
Portabilidad con versiones anteriores	HIVE-20618: Durante la unión, se BucketMap Join puede elegir la selección para tablas que no estén agrupadas
Portabilidad con versiones anteriores	HIVE-20617: corrección del tipo de constantes en las expresiones IN para que tengan el tipo correcto
Portabilidad con versiones anteriores	HIVE-20612: creación de un nuevo indicador de correlación multiclave de unión para la CBO

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20603: error “FS incorrecto” al insertar en partición después de cambiar la ubicación de la tabla o el sistema de archivos
Portabilidad con versiones anteriores	HIVE-20601: nulo en el evento ALTER_PARTITION en EnvironmentContext DbNotificationListener
Portabilidad con versiones anteriores	HIVE-20583: Utilice el nombre de host canónico solo para la autenticación de kerberos en HiveConnection
Portabilidad con versiones anteriores	HIVE-20582: conversión de hflush en el registro de protocolos de Hive a configurable
Portabilidad con versiones anteriores	HIVE-20563: vectorización: la expresión CASE WHEN falla cuando el tipo THEN/ELSE y el tipo del resultado son diferentes
Portabilidad con versiones anteriores	HIVE-20558: cambio del valor predeterminado de hive.hashtable.key.count.adjustment a 0,99
Portabilidad con versiones anteriores	HIVE-20552: Obtenga LogicalPlan un esquema más rápido
Portabilidad con versiones anteriores	HIVE-20550: Switch Web HCat para usar Beeline para enviar consultas de Hive
Portabilidad con versiones anteriores	HIVE-20537: estimaciones de unión de varias columnas con columnas no correlacionadas diferentes en CBO y Hive
Portabilidad con versiones anteriores	HIVE-20524: la comprobación de la evolución del esquema se interrumpe al pasar de la versión 2 a la versión 3 de Hive para ALTER TABLE VARCHAR a DECIMAL

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20522: HiveFilterSetOpTransposeRule puede generar un error de afirmación debido a la nulabilidad de los campos
Portabilidad con versiones anteriores	HIVE-20521: HS2 doas=True tiene problemas de permisos con hadoop.tmp.dir, con los sistemas de archivos MR y S3A
Portabilidad con versiones anteriores	HIVE-20515: los resultados de la consulta están vacíos cuando se utiliza la caché de resultados y el directorio temporal de consultas , el directorio de la caché de resultados en diferentes sistemas de archivos
Portabilidad con versiones anteriores	HIVE-20508: Hive no admite nombres de usuario del tipo “user@realm”
Portabilidad con versiones anteriores	HIVE-20507: Beeline: adición de un comando de utilidad para recuperar todos los URI de beeline-site.xml
Portabilidad con versiones anteriores	HIVE-20505: actualización de org.openj dk.jmh:jmh-core a 1.21
Portabilidad con versiones anteriores	HIVE-20503: uso de las estimaciones basadas en la estructura de datos durante la selección de mapjoin
Portabilidad con versiones anteriores	HIVE-20498: tipo de fecha de compatibilidad para la recopilación automática de estadísticas de columnas
Portabilidad con versiones anteriores	HIVE-20496: Vectorización: PTF vectorizado IllegalStateException

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20494: El genérico no funciona UDFRestrict InformationSchema después de HIVE-19440
Portabilidad con versiones anteriores	HIVE-20477: no se muestra si la expresión contiene OptimizedSql INs
Portabilidad con versiones anteriores	HIVE-20467: Permitir si no EXISTS/IF EXISTS in Resource plan creation/drop
Portabilidad con versiones anteriores	HIVE-20462: “CREATE VIEW IF NOT EXISTS” falla si la vista ya existe
Portabilidad con versiones anteriores	HIVE-20455: Velocidad de registro extraída de security.authorization. PrivilegeSynchroniz er.ejecutar
Portabilidad con versiones anteriores	HIVE-20439: uso del límite de memoria inflado al seleccionar la unión para Llap
Portabilidad con versiones anteriores	HIVE-20433: la conversión implícita de cadena a marca de tiempo es lenta
Portabilidad con versiones anteriores	HIVE-20432: se reescribe BETWEEN a IN para tipos enteros para estimación de estadísticas
Portabilidad con versiones anteriores	HIVE-20423: se establece NULLS LAST como el orden nulo predeterminado
Portabilidad con versiones anteriores	HIVE-20418: es posible que LLAP IO no gestione los archivos ORC que tienen el índice de filas desactivado correctamente para consultas sin columnas seleccionadas
Portabilidad con versiones anteriores	HIVE-20412: Entrada NPE HiveMetaHook

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20406: Coalesce devuelve arroja resultados incorrectos
Portabilidad con versiones anteriores	HIVE-20399: CTAS con una ubicación de tabla personalizada que no esté completamente calificada fallan en las tablas MM
Portabilidad con versiones anteriores	HIVE-20393: Reducción de semiuniones: el dPP se comporta de forma incoherente markSemiJoin
Portabilidad con versiones anteriores	HIVE-20391: puede deducir un tipo de retorno incorrecto al descomponer la función agregada HiveAggregateReduceFunctionsRule
Portabilidad con versiones anteriores	HIVE-20383: el nombre de la cola no es válido y hay problemas de sincronización en el enlace de eventos de Hive Proto.
Portabilidad con versiones anteriores	HIVE-20367: vectorización: compatibilidad de streaming para PTF AVG, MAX, MIN, SUM
Portabilidad con versiones anteriores	HIVE-20366: las estimaciones de estado de la consulta TPC-DS 78 están desactivadas porque su filtro es nulo
Portabilidad con versiones anteriores	HIVE-20364: actualización predeterminada para hive.map.aggr.hash.min.reduction
Portabilidad con versiones anteriores	HIVE-20352: vectorización: compatibilidad de función de agrupamiento
Portabilidad con versiones anteriores	HIVE-20347: hive.optimize.sort.dynamic. partition debería funcionar con CTAS y MV particionados

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20345: la eliminación de la base de datos puede bloquearse si las tablas se eliminan de una llamada diferente
Portabilidad con versiones anteriores	HIVE-20343: Hive 3: CTAS no respeta transacti onal_properties
Portabilidad con versiones anteriores	HIVE-20340: El druida necesita que se explicite la marca de tiempo a la cadena cuando la salida CASTs de la función de marca de tiempo se utiliza como cadena
Portabilidad con versiones anteriores	HIVE-20339: vectorización: eliminación de una restricción innecesaria que provoca que algunos PTF con RANK no se vectoricen
Portabilidad con versiones anteriores	HIVE-20337:: Expr no está rellenando la lista de particiones correctamente CachedStore getPartitionsBy
Portabilidad con versiones anteriores	HIVE-20336: políticas de enmascaramiento y filtrado para vistas materializadas
Portabilidad con versiones anteriores	HIVE-20326: creación de restricciones con RELY como predeterminado en lugar de NO RELY
Portabilidad con versiones anteriores	HIVE-20321: Vectorización: reduzca el tamaño de la memoria de 1 columna a <1 VectorHas hKeyWrapper CacheLine
Portabilidad con versiones anteriores	HIVE-20320: activa la marca hive.opti mize.remove.sq_count_check

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20315: vectorización: se corrigen más problemas con los resultados nulos o incorrect os y se evitan conversiones o emisiones innecesarias
Portabilidad con versiones anteriores	HIVE-20314: incluye la reducción de particion es en la reescritura de vistas materializadas
Portabilidad con versiones anteriores	HIVE-20312: Permite que los clientes Arrow usen los suyos propios con BufferAllocator LlapOutputFormatService
Portabilidad con versiones anteriores	HIVE-20302: LLAP: la ejecución no vectoriza da en IO ignora las columnas virtuales, incluida ROW__ID
Portabilidad con versiones anteriores	HIVE-20300: VectorFileSinkArrowOperator
Portabilidad con versiones anteriores	HIVE-20299: posible carrera en la prueba unitaria de firmante de LLAP
Portabilidad con versiones anteriores	HIVE-20296: Mejore HivePointLookupOpt imizerRule para poder extraer de contextos más sofisticados
Portabilidad con versiones anteriores	HIVE-20294: vectorización: se corrigen problemas de resultados nulos o incorrectos en COALESCE/ELT
Portabilidad con versiones anteriores	HIVE-20292: orden de unión incorrecto en query93 tpcds con la restricción principal definida
Portabilidad con versiones anteriores	HIVE-20290: Inicialización lenta para no asignar ArrowColumnarBatchSerDe búferes durante GetSplits

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20281: SharedWorkOptimizer falla porque «el operador almacena en caché el contenido y el plan real son diferentes»
Portabilidad con versiones anteriores	HIVE-20277: vectorización: expresiones de casos que devuelven BOOLEAN no compatibl es con FILTER
Portabilidad con versiones anteriores	HIVE-20267: ampliación de la WebUI para incluir un formulario para configurar dinámicam ente los niveles de registro
Portabilidad con versiones anteriores	HIVE-20263: Error tipográfico en la variable HiveReduceExpressionsWithStatsRule
Portabilidad con versiones anteriores	HIVE-20260: NDV de una columna no debe escalarse cuando el filtro de otra columna cambia el recuento de filas
Portabilidad con versiones anteriores	HIVE-20252: reducción de semiuniones: los ciclos debidos a la ramificación de la semiunión pueden pasar desapercibidos si un lado pequeño de la tabla tiene una unión de mapas ascendente.
Portabilidad con versiones anteriores	HIVE-20245: vectorización: se corrigen problemas de resultados NULL o incorrectos en BETWEEN o IN
Portabilidad con versiones anteriores	HIVE-20241: compatibilidad con especificación de partición en instrucciones CTAS
Portabilidad con versiones anteriores	HIVE-20240: reducción de semiuniones: uso de variable local para comprobar el estado de la tabla externa

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20226: HMS getNextNotification generará una excepción cuando los maxEvents de la solicitud superen las filas máximas de la tabla
Portabilidad con versiones anteriores	HIVE-20225: para admitir el formato binario de Teradata SerDe
Portabilidad con versiones anteriores	HIVE-20213: actualización de Calcite a 1.17.0
Portabilidad con versiones anteriores	HIVE-20212: Hiveserver2 en modo HTTP emite incorrectamente la métrica default.General.open_connections
Portabilidad con versiones anteriores	HIVE-20210: El optimizador de recuperación simple debería permitir filtrar en una columna que no sea de partición y la conversión sea mínima MapReduce
Portabilidad con versiones anteriores	HIVE-20209: la conexión con el metaalmacén falla en el primer intento en el volcado de respuesta
Portabilidad con versiones anteriores	HIVE-20207: vectorización: se corrigen problemas de resultados nulos o incorrectos al filtrar o comparar
Portabilidad con versiones anteriores	HIVE-20204: conversión de tipos durante IN
Portabilidad con versiones anteriores	SerDe HIVE-20203: Arrow filtra un DirectByteBuffer
Portabilidad con versiones anteriores	HIVE-20197: Vectorización: añada las pruebas DECIMAL_64, añada la aritmética y añada más GRUPOS POR AGREGACIÓN Date/Interval/Timestamp

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20193: cboInfo no está presente en el JSON del plan explicativo
Portabilidad con versiones anteriores	HS2 HIVE-20192: con un JDOPersistence metastore integrado se filtran objetos de Manager
Portabilidad con versiones anteriores	HIVE-20183: la inserción desde una tabla con buckets puede provocar la pérdida de datos si la tabla de origen contiene un bucket vacío
Portabilidad con versiones anteriores	HIVE-20177: Vectorización: reduce la asignació n en el modo Streaming KeyWrapper GroupBy
Portabilidad con versiones anteriores	HIVE-20174: vectorización: se corrigen problemas de resultados nulos o incorrectos en las funciones de agregación GROUP BY
Portabilidad con versiones anteriores	HIVE-20172: se produjo un StatsUpdater error con una excepción GSS al intentar conectarse a un metabastore remoto
Portabilidad con versiones anteriores	HIVE-20153: las UDF de recuento y suma consumen más memoria en Hive 2+
Portabilidad con versiones anteriores	HIVE-20152: restablecimiento del estado de la base de datos, cuando repl dump falla, por lo que se puede cambiar el nombre de la tabla
Portabilidad con versiones anteriores	HIVE-20149: TestHiveCli error o tiempo de espera
Portabilidad con versiones anteriores	HIVE-20130: mejor registro para el sincroniz ador de esquemas de información

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20129: reversión a la evolución del esquema basado en la posición para las tablas orc
Portabilidad con versiones anteriores	SessionStateUserAuthenticatorHIVE-2018:. getGroupNames
Portabilidad con versiones anteriores	HIVE-20116: TezTask está utilizando el registrador principal
Portabilidad con versiones anteriores	HIVE-20115: las tablas de Acid no deben usar análisis de pie de página
Portabilidad con versiones anteriores	HIVE-20103: WM: solo se agregan contadores de DAG si se utiliza al menos uno
Portabilidad con versiones anteriores	HIVE-20101: BloomKFilter: evite utilizar por completo las matrices de bytes [] locales
Portabilidad con versiones anteriores	HIVE-20100: OpTraits : Select Optraits debería detenerse cuando se detecte una discordancia
Portabilidad con versiones anteriores	HIVE-20098: estadísticas: NPE al obtener las estadísticas de partición de la columna de fecha
Portabilidad con versiones anteriores	HIVE-20095: corrección de la característica para enviar el cálculo a las tablas externas de JDBC
Portabilidad con versiones anteriores	HIVE-20093: Se usa con Netty para la contabili dad LlapOutputFomatService ArrowBuf
Portabilidad con versiones anteriores	HIVE-20090: ampliación de la creación de filtros reductores de semiuniones para poder descubrir nuevas oportunidades

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20088: la ruta de ubicación de configura ción de Beeline está ensamblada incorrect amente
Portabilidad con versiones anteriores	HIVE-20082: la conversión HiveDecimal a cadena no formatea el decimal correctamente
Portabilidad con versiones anteriores	HIVE-20069: corrección de la reoptimización en caso de optimización de DPP y Semijoin
Portabilidad con versiones anteriores	HIVE-20051: omisión de la autorización para tablas temporales
Portabilidad con versiones anteriores	HIVE-20044: Arrow Serde debería rellenar los valores de caracteres y gestionar correctam ente las cadenas vacías
Portabilidad con versiones anteriores	HIVE-20028: la configuración de la caché del cliente del metaalmacén se utiliza incorrect amente
Portabilidad con versiones anteriores	HIVE-20025: Limpieza de los archivos de eventos creados por HiveProtoLoggingHook
Portabilidad con versiones anteriores	HIVE-20020: el jar de contribuciones de Hive no debería estar en lib
Portabilidad con versiones anteriores	HIVE-20013: adición de una conversión implícita al tipo de fecha para la función to_date
Portabilidad con versiones anteriores	HIVE-20011: aléjese del modo de adición en el enlace de registro de prototipos
Portabilidad con versiones anteriores	HIVE-20005: acid_table_stats, acid_no_b uckets, etc.: el resultado de la consulta cambia en la rama

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-20004: Si 64 utiliza una escala incorrecta, se obtienen resultados incorrectos ConvertDecimal ToDecimal
Portabilidad con versiones anteriores	HIVE-19995: tráfico de filas agregado para tablas de Acid
Portabilidad con versiones anteriores	HIVE-19993: no es posible utilizar un alias de tabla que también aparezca como nombre de columna
Portabilidad con versiones anteriores	HIVE-19992: Vectorización: continuación de HIVE-19951 --> añadir llamada a. SchemaEvolution isOnlyImplicitConversión para deshabilitar las E/S LLAP codificadas para ORC solo cuando la conversión de tipos de datos no está implícita
Portabilidad con versiones anteriores	HIVE-19989: Metastore utiliza un nombre de aplicación incorrecto para las métricas HADOOP2
Portabilidad con versiones anteriores	HIVE-19981: Las tablas gestionadas convertidas en tablas externas por la HiveStrictManagedMigration utilidad deben configurarse para que eliminen los datos al eliminar la tabla
Portabilidad con versiones anteriores	HIVE-19967: SMB Join: Se necesitan retratos para una banda sonora PTFOperator
Portabilidad con versiones anteriores	HIVE-19935: sesión de Hive WM cerrada: no se pudo actualizar el recuento de tareas de LLAP
Portabilidad con versiones anteriores	HIVE-19924: etiquetado de los trabajos de distcp ejecutados por Repl Load

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-19891: la inserción en tablas externas con directorios de particiones personalizados puede provocar la pérdida de datos
Portabilidad con versiones anteriores	HIVE-19850: la poda dinámica de particiones en Tez provoca el error “No se encontró ningún trabajo para tablescan”
Portabilidad con versiones anteriores	HIVE-19806: ordenación de los resultados de las pruebas para evitar errores en los resultado s de las pruebas
Portabilidad con versiones anteriores	HIVE-19770: compatibilidad con CBO para consultas con varias columnas iguales al seleccionar
Portabilidad con versiones anteriores	HIVE-19769: creación de objetos dedicados para nombres de bases de datos y tablas
Portabilidad con versiones anteriores	HIVE-19765: Añada las pruebas específicas de Parquet a BlobstoreCliDriver
Portabilidad con versiones anteriores	HIVE-19759: Prueba escamosa: # TestRpc testServerPort
Portabilidad con versiones anteriores	HIVE-19711: herramienta de esquema de refactorización de Hive
Portabilidad con versiones anteriores	HIVE-19701: no necesita sincronizarse getDelegationToken FromMetaStore
Portabilidad con versiones anteriores	HIVE-19694: la instrucción Create Materiali zed View debe comprobar si hay conflictos de nombres de MV antes de ejecutar la instrucción SQL de MV.

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-19674: agrupación por constantes decimales hasta la tabla de Druid
Portabilidad con versiones anteriores	HIVE-19668: Más del 30% del montón se desperdicia al duplicar org.antlr.runtime. CommonTokens y cadenas duplicadas
Portabilidad con versiones anteriores	HIVE-19663: refactorización de la generación de informes LLAP IO
Portabilidad con versiones anteriores	HIVE-19661: cambia Hive UDFs para usar el motor de expresiones regulares Re2J
Portabilidad con versiones anteriores	HIVE-19628: posible NPE en LLAP TestSigning
Portabilidad con versiones anteriores	HIVE-19568: HA activa/pasiva: no permite la conexión directa a una instancia pasiva HS2 HS2
Portabilidad con versiones anteriores	HIVE-19564: vectorización: se corrigen problemas de resultados NULL o incorrectos en aritmética
Portabilidad con versiones anteriores	HIVE-19552 TestMiniDruidKafkaCliDriver: Habilitar #druidkafkamini_basic .q
Portabilidad con versiones anteriores	HIVE-19432: GetTablesOperation es demasiado lento si la colmena tiene demasiada s bases de datos y tablas
Portabilidad con versiones anteriores	HIVE-19360: CBO: Añadir un «SQL optimizad o» al objeto QueryPlan
Portabilidad con versiones anteriores	HIVE-19326: recopilación automática de estadísticas: agregación incorrecta durante las consultas UNION

Tipo	Descripción
Portabilidad con versiones anteriores	TestJdbcWithDBTokenStoreNoDoAsHIVE-19313: las pruebas están fallando
Portabilidad con versiones anteriores	HIVE-19285: Agregue registros a las subclases de MetaDataOperation
Portabilidad con versiones anteriores	HIVE-19235: actualización de los archivos dorados para las pruebas de Minimr
Portabilidad con versiones anteriores	HIVE-19104: Cuando se inicia la prueba con un reintento, las MetaStore instancias deben ser independientes
Portabilidad con versiones anteriores	HIVE-18986: El cambio de nombre de la tabla ejecutará java.lang. StackOverflowError en DataNucleus si la tabla contiene un gran número de columnas
Portabilidad con versiones anteriores	HIVE-18920: CBO: inicialización de los proveedores de Janino antes de la primera consulta
Portabilidad con versiones anteriores	HIVE-18873: Omitir la pulsación de predicado s para MR de forma silenciosa puede provocar que los controladores de almacenamiento produzcan resultados erróneos HiveInput Format
Portabilidad con versiones anteriores	HIVE-18871: error de ejecución de Hive en Tex porque hive.aux.jars.path está establecido en hdfs://
Portabilidad con versiones anteriores	HIVE-18725: mejora del manejo de errores en las subconsultas si hay una referencia de columna incorrecta

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-18696: Es posible que las carpetas de particiones no se limpien correctamente con el método .add_partitions_core si un HiveMetaStore
Portabilidad con versiones anteriores	HIVE-18453: ACID: agregación de la sintaxis “CREATE TRANSACTIONAL TABLE” para unificar la compatibilidad con ACID ORC y Parquet
Portabilidad con versiones anteriores	HIVE-18201: desactivación de XPROD_EDGE para sq_count_chec
Portabilidad con versiones anteriores	HIVE-18140: las estadísticas de las tablas particionadas pueden fallar en el caso de las estadísticas básicas
Portabilidad con versiones anteriores	HIVE-17921: la agregación con una estructura en LLAP produce un resultado incorrecto
Portabilidad con versiones anteriores	HIVE-17896: Parte superior: cree un operador Top vectorizable independiente NKey NKey
Portabilidad con versiones anteriores	HIVE-17840: elimina la excepción si TransactionalListeners.NotifyEvent falla HiveMetaStore
Portabilidad con versiones anteriores	HIVE-17043: eliminación de las columnas que no sean únicas del grupo por claves si no se hace referencia a ellas más adelante
Portabilidad con versiones anteriores	HIVE-17040: eliminación conjunta en presencia de una relación FK
Portabilidad con versiones anteriores	HIVE-16839: llamadas desequilibradas a openTransaction/commitTransaction cuando se altera la misma partición al mismo tiempo

Tipo	Descripción
Portabilidad con versiones anteriores	HIVE-16100: el optimizador dinámico de particiones clasificadas pierde operadores secundarios
Portabilidad con versiones anteriores	StackOverflowError HIVE-15956: cuando se descartan muchas particiones
Portabilidad con versiones anteriores	HIVE-15177: la autenticación con Hive falla cuando el tipo de autenticación de Kerberos está establecido en fromSubject y la entidad principal contiene _HOST
Portabilidad con versiones anteriores	HIVE-14898: no HS2 debería registrar la pila de llamadas por un error en el encabezado de autenticación vacío
Portabilidad con versiones anteriores	HIVE-14493: compatibilidad de particion amiento con vistas materializadas
Portabilidad con versiones anteriores	HIVE-14431: reconocimiento de COALESCE como CASE
Portabilidad con versiones anteriores	HIVE-13457: Cree puntos finales de la API REST para monitorear la información HS2
Portabilidad con versiones anteriores	HIVE-12342: establecimiento del valor predeterminado de hive.optimize.index.filter en verdadero
Portabilidad con versiones anteriores	HIVE-10296: se observa una excepción de conversión cuando Hive ejecuta una consulta de unión múltiple en el metaalmacén
Portabilidad con versiones anteriores	HIVE-6980: eliminación de la tabla mediante SQL directo

Amazon EMR 6.6.0: cambios de configuración en Hive

- Como parte del cambio de HIVE-20703 de OSS, se sustituyó la propiedad para ordenar las particiones dinámicas, `hive.optimize.sort.dynamic.partition`, por `hive.optimize.sort.dynamic.partition.threshold`.

  La configuración de `hive.optimize.sort.dynamic.partition.threshold` tiene los siguientes valores potenciales:

Valor	Descripción
`0` (predeterminado)	Hace que la optimización para ordenar las particiones dinámicas sea una decisión basada en los costos cuando se utilizan archivos ORC. El número máximo de escritores permitidos en las consultas INSERT se calcula dividiendo (memoria del ejecutor/contenedor) * (porcentaje de memoria ocupado por orc) por el máximo de memoria (tamaño de banda) ocupado por un solo escritor.
`-1`	Desactiva la optimización para ordenar completamente las particiones dinámicas.
`1`	Permite la clasificación global de las particiones dinámicas. Esto mantiene abierto solo un escritor de registros para cada valor de partición del reductor, lo que reduce la presión de memoria sobre los reductores.
`2` (o un entero mayor)	Indica a Hive que utilice el entero especificado como umbral para el número máximo de escritores.

Amazon EMR 6.6.0: problemas conocidos de Hive

- Las consultas con funciones de ventana en la misma columna que la de unión pueden provocar transformaciones no válidas, como se indica en HIVE-25278, y provocar resultados incorrectos o errores en las consultas. Como solución alternativa, puede deshabilitar la CBO en el nivel de consulta para este tipo de consultas. Póngase en contacto con AWS el servicio de asistencia para obtener más información.

- Amazon EMR 6.6.0 incluye la versión 3.1.2 del software Hive. Hive 3.1.2 presenta una característica que divide los archivos de texto si contienen un encabezado y un pie de página (HIVE-21924). Apache Tez App Master lee cada uno de sus archivos para determinar los puntos de desfase en el rango de datos. La combinación de estos comportamientos podría afectar negativamente al rendimiento si sus consultas leen una gran cantidad de archivos de texto pequeños. Como solución alternativa, utilice `CombineHiveInputFormat` y ajuste el tamaño máximo de división configurando las siguientes propiedades:

```
SET hive.tez.input.format=org.apache.hadoop.hive.ql.io.CombineHiveInputFormat;
SET mapreduce.input.fileinputformat.split.maxsize=16777216;
```

- Con las versiones 6.6.0 a 6.9.x de Amazon EMR, las consultas INSERT con partición dinámica y una cláusula ORDER BY o SORT BY siempre tendrán dos reductores. Este problema se debe al cambio HIVE-20703 de OSS, que sitúa la optimización de la clasificación dinámica de las particiones en una decisión basada en los costos. Si su carga de trabajo no requiere ordenar las particiones dinámicas, le recomendamos que establezca la propiedad `hive.optimize.sort.dynamic.partition.threshold` en `-1` para deshabilitar la nueva característica y obtener el número de reductores calculado correctamente. Este problema se ha corregido en OSS Hive como parte de HIVE-22269 y se ha corregido en Amazon EMR 6.10.0.

Hudi

Apache Hudi es un marco de administración de datos de código abierto que se utiliza para simplificar el procesamiento incremental de los datos y el desarrollo de canalizaciones de datos, porque proporciona, en el nivel de registro, prestaciones de inserción, actualización, upsert y eliminación. Upsert se refiere a la capacidad de insertar registros en un dataset si aún no existen o de actualizarlos en caso de que ya estén presentes. Al administrar de manera eficiente la forma en que se establecen los datos en Amazon S3, Hudi permite que los datos se inserten y actualicen prácticamente en tiempo real. Hudi mantiene cuidadosamente los metadatos de las acciones realizadas en el conjunto de datos para ayudar a garantizar que las acciones sean atómicas y coherentes.

Hudi está integrado con Apache Spark, Apache Hive y Presto. En las versiones 6.1.0 y posteriores de Amazon EMR, Hudi también está integrado con Trino (PrestoSQL).

Con la versión 5.28.0 y posteriores de Amazon EMR, EMR instala los componentes de Hudi de forma predeterminada cuando se instalan Spark, Hive, Presto o Flink. Puedes usar Spark o la DeltaStreamer utilidad Hudi para crear o actualizar conjuntos de datos Hudi. Puede utilizar Hive, Spark, Presto o Flink para consultar un conjunto de datos de Hudi de forma interactiva o crear canalizaciones de procesamiento de datos mediante la extracción incremental. La extracción incremental se refiere a la capacidad de extraer únicamente los datos que han cambiado entre dos acciones.

Estas características hacen que Hudi sea adecuado para los siguientes casos de uso:

- Trabajo con datos de transmisión desde sensores y otros dispositivos del Internet de las cosas (IoT) que requieren eventos específicos de inserción y actualización de datos.
- Cumplimiento de la normativa de privacidad de datos en aplicaciones cuyos usuarios pueden ejercer su derecho al olvido o modificar su consentimiento para el uso de sus datos.
- Implementación de un sistema de captura de datos de cambio (CDC) que permita aplicar cambios a un dataset a lo largo del tiempo.

En la siguiente tabla, se muestra la versión de Hudi incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Hudi.

Para ver la versión de los componentes instalados con Hudi en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Hudi para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Hudi	Componentes instalados con Hudi
emr-7.9.0	Hudi 0.15.0-amzn-6	Not available.

En la tabla siguiente, se muestra la versión de Hudi incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Hudi.

Para ver la versión de los componentes instalados con Hudi en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Hudi para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Hudi	Componentes instalados con Hudi
emr-6.15.0	Hudi 0.14.0-amzn-0	Not available.

> **Note**
>
> Amazon EMR 6.8.0 viene con Apache Hudi 0.11.1; sin embargo, los clústeres de Amazon EMR 6.8.0 también son compatibles con el código abierto `hudi-spark3.3-bundle_2.12` de Hudi 0.12.0.

En la tabla siguiente, se muestra la versión de Hudi incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Hudi.

Para ver la versión de los componentes instalados con Hudi en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Hudi para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Hudi	Componentes instalados con Hudi
emr-5.36.2	Hudi 0.10.1-amzn-1	Not available.

Temas

- Cómo funciona Hudi
- Consideraciones y limitaciones para usar Hudi en Amazon EMR
- Crear un clúster con Hudi instalado
- Uso de un conjunto de datos de Hudi
- Usar la CLI de Hudi
- Historial de versiones de Hudi

Cómo funciona Hudi

Al utilizar Hudi con Amazon EMR, puede escribir datos en el conjunto de datos mediante la API de origen de datos de Spark o la DeltaStreamer utilidad Hudi. Hudi organiza un conjunto de datos en una estructura de directorios particionada en una estructura similar a la de *`basepath`*, que es similar a una tabla de Hive tradicional. Los detalles de cómo se disponen los datos en archivos en estos directorios dependen del tipo de dataset que se elija. Puede elegir entre copiar al escribir (CoW, Copy on Write) o fusionar al leer (MOR, Merge on Read).

Independientemente del tipo de dataset, cada partición que contiene se identifica de forma única por su `partitionpath` relativa a `basepath`. Dentro de cada partición, los registros se distribuyen en varios archivos de datos. Para obtener más información, consulte el artículo sobre administración de archivos en la documentación de Apache Hudi.

Cada acción de Hudi tiene una confirmación correspondiente, que se identifica mediante una marca de tiempo que aumenta de forma monótona y se conoce como Instant. Hudi guarda una serie de todas las acciones hechas en el conjunto de datos en forma de cronología. Hudi se basa en la cronología para aislar instantáneamente a lectores y escritores, y para poder retroceder a un punto anterior en el tiempo. Para obtener más información acerca de las acciones que Hudi registra y el estado de las acciones, consulte Cronologías en la documentación de Apache Hudi.

Descripción de los tipos de almacenamiento de conjuntos de datos: escribir frente a combinar al leer

Cuando se crea un conjunto de datos de Hudi, se especifica si será de tipo copiar al escribir o fusionar al leer.

- Copiar al escribir (CoW, Copy on Write): los datos se almacenan en un formato de columnas (Parquet) y cada actualización crea una nueva versión de los archivos durante una escritura. CoW es el tipo de almacenamiento predeterminado.
- Fusionar al leer (MoR, Merge on Read): los datos se almacenan utilizando un formato que combina columnas (Parquet) y filas (Avro). Las actualizaciones se registran en archivos delta basados en filas y se compactan según sea necesario para crear nuevas versiones de los archivos en columnas.

Con los datasets de tipo CoW, cada vez que se produce una actualización de un registro, el archivo que contiene el registro se vuelve a escribir con los valores actualizados. Con un conjunto de datos de tipo MoR, cada vez que hay una actualización, Hudi escribe solo la fila correspondiente al registro modificado. MoR es más adecuado para cargas de trabajo con gran cantidad de escrituras o cambios y menos lecturas. CoW es más adecuado para cargas de trabajo con gran cantidad de lecturas con datos que cambian con menos frecuencia.

Hudi proporciona tres vistas lógicas para acceder a los datos:

- Vista optimizada de lectura: proporciona el último conjunto de datos confirmado a partir de las tablas CoW y el último conjunto de datos compactado a partir de las tablas MoR.
- Vista incremental: proporciona una secuencia de cambios entre dos acciones procedentes de un conjunto de datos CoW para alimentar los trabajos posteriores y extraer, transformar y los flujos de trabajo de extracción, transformación y carga (ETL).
- Vista en tiempo real: proporciona los datos confirmados más recientes de una tabla MoR; para ello, fusiona en línea los archivos en columnas y basados en filas.

Cuando se consulta la vista optimizada para lectura, la consulta devuelve todos los datos compactados, pero no incluye las últimas confirmaciones delta. La consulta de estos datos proporciona un buen rendimiento de lectura, pero omite los datos más recientes. Cuando se consulta la vista en tiempo real, Hudi fusiona los datos compactados con las confirmaciones delta durante la lectura. Los datos más recientes están disponibles para consultarlos, pero fusionarlos genera una sobrecarga computacional que hace que la consulta sea menos eficiente. La capacidad de consultar datos compactados o datos en tiempo real permite elegir entre rendimiento y flexibilidad al realizar consultas.

Para obtener más información acerca de las ventajas e inconvenientes de ambos los tipos de almacenamiento, consulte Tipos y vistas de almacenamiento en la documentación de Apache Hudi.

Hudi crea dos tablas en el metaalmacén de Hive para MoR: una tabla con el nombre que especificó, que es una vista optimizada para lectura, y una tabla con el mismo nombre al que se ha anexado `_rt`, que es una vista en tiempo real. Puede consultar ambas tablas.

Registro de un conjunto de datos de Hudi con su metaalmacén

Si registra una tabla de Hudi con el metaalmacén de Hive, puede consultar las tablas de Hudi mediante Hive, Spark SQL o Presto como haría con cualquier otra tabla. Además, puede integrar Hudi con AWS Glue si configura Hive y Spark para usar el Catálogo de datos de AWS Glue como metaalmacén. Para las tablas de MoR, Hudi registra el conjunto de datos como dos tablas en el metaalmacén: una tabla con el nombre que especificó, que es una vista optimizada para lectura, y una tabla con el mismo nombre al que se ha anexado `_rt`, que es una vista en tiempo real.

Para registrar una tabla de Hudi con el metaalmacén de Hive al usar Spark para crear un conjunto de datos de Hudi, configure la opción `HIVE_SYNC_ENABLED_OPT_KEY` en `"true"` e indique otras propiedades requeridas. Para obtener más información, consulte Uso de un conjunto de datos de Hudi. Además, puede utilizar la utilidad de línea de comandos hive_sync_tool para registrar un conjunto de datos de Hudi como tabla en el metaalmacén, por separado.

Consideraciones y limitaciones para usar Hudi en Amazon EMR

- El campo de clave de registro no puede ser nulo ni estar vacío: el campo especificado como campo de clave de registro no puede tener valores `null` ni estar vacío.
- Esquema actualizado de forma predeterminada en upsert e insert: Hudi proporciona una interfaz `HoodieRecordPayload` que determina cómo se fusionan el conjunto de datos Hudi existente DataFrame y de entrada para producir un conjunto de datos nuevo y actualizado. Hudi proporciona una implementación predeterminada de esta clase`OverwriteWithLatestAvroPayload`, que sobrescribe los registros existentes y actualiza el esquema como se especifica en la entrada. DataFrame Para personalizar esta lógica con el fin de implementar actualizaciones parciales y fusiones, puede proporcionar una implementación de la interfaz `HoodieRecordPayload` mediante el parámetro `DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY`.
- La eliminación requiere un esquema: al eliminar, debe especificar la clave de registro, la clave de partición y los campos de clave de combinación previa. Se puede hacer que otras columnas sean `null` o estén vacías, pero se requiere el esquema completo.
- Limitaciones de las tablas MoR: las tablas MoR no admiten puntos de guardado. Puede consultar las tablas MoR mediante la vista optimizada para lectura o la vista en tiempo real (*`tableName`*`_rt`) de Spark SQL, Presto o Hive. El uso de la vista optimizada para lectura solo

expone los datos del archivo base y no expone una vista combinada de los datos base y de registro.

- Hive
  - Para registrar tablas en el metaalmacén de Hive, Hudi espera que el servidor Hive Thrift se ejecute en el puerto predeterminado, `10000`. Si se sustituye este puerto por un puerto personalizado, es preciso pasar la opción `HIVE_URL_OPT_KEY` como se muestra en el siguiente ejemplo.

    ```
    .option(DataSourceWriteOptions.HIVE_URL_OPT_KEY, "jdbc:hive2://localhost:override-port-number
    ```

  - El tipo de datos `timestamp` de Spark se registra como tipo de datos `long` en Hive, pero no como el tipo `timestamp` de Hive.
- Presto
  - Presto no admite la lectura de tablas MoR en tiempo real en versiones de Hudi anteriores a 0.6.0.
  - Presto solo admite consultas de instantáneas.
  - Para que Presto interprete correctamente las columnas del conjunto de datos de Hudi, establezca el valor de `hive.parquet_use_column_names` en `true`.
    - Para establecer el valor de una sesión, en el shell de Presto, ejecute el siguiente comando:

      ```
      set session hive.parquet_use_column_names=true
      ```

    - Para establecer el valor en el nivel de clúster, utilice la clasificación de configuración `presto-connector-hive` para establecer `hive.parquet.use_column_names` en `true`, como se muestra en el ejemplo siguiente. Para obtener más información, consulte Configuración de aplicaciones.

      ```
      [
        {
          "Classification": "presto-connector-hive",
          "Properties": {
            "hive.parquet.use-column-names": "true"
          }
        }
      ]
      ```

- HBase Índice

- La HBase versión utilizada para compilar Hudi puede ser diferente de la que aparece en la Guía de versiones de EMR. Para obtener las dependencias correctas para la sesión de Spark, ejecute el siguiente comando.

```
spark-shell \
--jars /usr/lib/spark/external/lib/spark-avro.jar,/usr/lib/hudi/cli/lib/*.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false"
```

- Configuración para obtener el mejor rendimiento: para EMR 7.3+/Hudi 0.15+, se recomienda a los clientes establecer esta configuración para reducir la sobrecarga de serialización de Kryo:

```
--conf 'spark.kryo.registrator=org.apache.spark.HoodieKryoRegistrar'
```

> **Note**
>
> Si utiliza un control de acceso detallado (FGAC) en EMR Serverless, esta configuración no es necesaria, ya que los usuarios deben utilizarla en lugar de. JavaSerializer KryoSerializer

Crear un clúster con Hudi instalado

Con la versión 5.28.0 y posteriores de Amazon EMR, Amazon EMR instala los componentes de Hudi de forma predeterminada cuando se instalan Spark, Hive o Presto. Para utilizar Hudi en Amazon EMR, cree un clúster con una o varias de las siguientes aplicaciones instaladas:

- Hadoop
- Hive
- Spark
- Presto
- Flink

Puede crear un clúster mediante la AWS Management Console AWS CLI, la o la API de Amazon EMR.

Para crear un clúster con Hudi mediante la AWS Management Console

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. En Software Configuration (Configuración de software), elija emr-5.28.0 o posterior para Release (Versión) y seleccione Hadoop, Hive, Spark, Presto y Tez, así como las demás aplicaciones que el clúster requiera.
4. Configure otras opciones según sea necesario para la aplicación y, a continuación, elija Next (Siguiente).
5. Configure las opciones de Hardware (Hardware) y General cluster settings (Configuración general del clúster) según lo desee.
6. Para las opciones de seguridad, le recomendamos que seleccione un EC2 key pair que pueda usar para conectarse a la línea de comandos del nodo maestro mediante SSH. Esto le permitirá ejecutar los comandos del intérprete de comandos de Spark, los comandos de la CLI de Hudi y los comandos de la CLI de Hudi descritos en esta guía.
7. Configure las demás opciones de seguridad como desee y elija Create cluster (Crear clúster).

Uso de un conjunto de datos de Hudi

Hudi permite insertar, actualizar y eliminar datos de conjuntos de datos de Hudi a través de Spark. Para obtener más información, consulte Escritura de tablas de Hudi en la documentación de Apache Hudi.

En los siguientes ejemplos, se muestra cómo iniciar el intérprete de comandos interactivo de Spark, utilizar el envío de Spark o usar Cuadernos de Amazon EMR para trabajar con Hudi en Amazon EMR. También puede utilizar la DeltaStreamer utilidad Hudi u otras herramientas para escribir en un conjunto de datos. En esta sección, los ejemplos muestran cómo trabajar con datasets usando el shell de Spark mientras está conectado al nodo principal usando SSH como usuario predeterminado de `hadoop`.

Iniciar el intérprete de comandos de Spark con Amazon EMR 6.7 y versiones posteriores

Cuando ejecute `spark-shell`, `spark-submit` o `spark-sql` con Amazon EMR 6.7.0 o una versión posterior, ejecute los siguientes comandos.

> **Note**
>
> Amazon EMR 6.7.0 usa Apache Hudi 0.11.0-amzn-0, que contiene mejoras significativas con respecto a las versiones anteriores de Hudi. Para obtener más información, consulte la Guía de migración a Apache Hudi 0.11.0. Los ejemplos de esta pestaña reflejan estos cambios.

Para abrir el intérprete de comandos de Spark en el nodo principal

1. Conéctese al nodo principal mediante SSH. Para obtener más información, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR.
2. Introduzca el siguiente comando para iniciar el shell de Spark. Para usar la PySpark carcasa, sustitúyala por*spark-shell*. *pyspark*

```
spark-shell --jars /usr/lib/hudi/hudi-spark-bundle.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.hudi.catalog.HoodieCatalog"
 \
--conf "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
```

Iniciar el intérprete de comandos de Spark con Amazon EMR 6.6 y versiones anteriores

Cuando ejecute `spark-shell`, `spark-submit` o `spark-sql` con Amazon EMR 6.6.x o una versión posterior, ejecute los siguientes comandos.

> **Note**
>
> - Amazon EMR 6.2 y 5.31 y versiones posteriores (Hudi 0.6.x y versiones posteriores) pueden omitir `spark-avro.jar` en la configuración.

- Amazon EMR 6.5 y 5.35 y versiones posteriores (Hudi 0.9.x y versiones posteriores) pueden omitir `spark.sql.hive.convertMetastoreParquet=false` en la configuración.
- Amazon EMR 6.6 y 5.36 y versiones posteriores (Hudi 0.10.x y versiones posteriores) deben incluir la configuración de `HoodieSparkSessionExtension` descrita en la Guía de la versión 0.10.0 de Spark:

```
--conf
 "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
 \
```

Para abrir el intérprete de comandos de Spark en el nodo principal

1. Conéctese al nodo principal mediante SSH. Para obtener más información, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR.
2. Introduzca el siguiente comando para iniciar el shell de Spark. Para usar la PySpark carcasa, *`spark-shell`* sustitúyala por*`pyspark`*.

```
spark-shell \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false" \
--jars /usr/lib/hudi/hudi-spark-bundle.jar,/usr/lib/spark/external/lib/spark-
avro.jar
```

Utilizar Hudi con Cuadernos de Amazon EMR que utilicen Amazon EMR 6.7 y versiones posteriores

Para utilizar Hudi con Cuadernos de Amazon EMR, antes debe copiar los archivos jar de Hudi del sistema de archivos local a HDFS en el nodo maestro del clúster de bloc de notas. A continuación, puede utilizar el editor de cuadernos para configurar el cuaderno de EMR de modo que utilice Hudi.

Para usar Hudi con Cuadernos de Amazon EMR

1. Cree y lance un clúster para Cuadernos de Amazon EMR. Para más información, consulte Creación de clústeres de Amazon EMR para cuadernos en la Guía de administración de Amazon EMR.

2. Conéctese al nodo principal del clúster mediante SSH y, a continuación, copie los archivos jar desde el sistema de archivos local a HDFS, como se muestra en los ejemplos siguientes. En el ejemplo, creamos un directorio en HDFS para que la administración de archivos resulte más clara. Puede elegir su propio destino en HDFS, si lo desea.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-spark-bundle.jar
```

3. Abra el editor de bloc de notas, escriba el código del siguiente ejemplo y ejecútelo.

```
%%configure
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar",
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.hudi.catalog.HoodieCatalog",

 "spark.sql.extensions":"org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
          }}
```

Utilizar Hudi con Cuadernos de Amazon EMR que utilicen Amazon EMR 6.6 y versiones anteriores

Para utilizar Hudi con Cuadernos de Amazon EMR, antes debe copiar los archivos jar de Hudi del sistema de archivos local a HDFS en el nodo maestro del clúster de bloc de notas. A continuación, puede utilizar el editor de cuadernos para configurar el cuaderno de EMR de modo que utilice Hudi.

Para usar Hudi con Cuadernos de Amazon EMR

1. Cree y lance un clúster para Cuadernos de Amazon EMR. Para más información, consulte Creación de clústeres de Amazon EMR para cuadernos en la Guía de administración de Amazon EMR.
2. Conéctese al nodo principal del clúster mediante SSH y, a continuación, copie los archivos jar desde el sistema de archivos local a HDFS, como se muestra en los ejemplos siguientes. En el ejemplo, creamos un directorio en HDFS para que la administración de archivos resulte más clara. Puede elegir su propio destino en HDFS, si lo desea.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-
spark-bundle.jar
```

```
hdfs dfs -copyFromLocal /usr/lib/spark/external/lib/spark-avro.jar /apps/hudi/lib/
spark-avro.jar
```

3. Abra el editor de bloc de notas, escriba el código del siguiente ejemplo y ejecútelo.

```
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar,hdfs:///apps/
hudi/lib/spark-avro.jar",
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.hive.convertMetastoreParquet":"false"
          }}
```

Iniciar una sesión de Spark para Hudi

Al usar Scala, debe importar las siguientes clases a su sesión de Spark. Esto debe hacerse una vez por cada sesión de Spark.

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._
import org.apache.hudi.DataSourceWriteOptions
import org.apache.hudi.DataSourceReadOptions
import org.apache.hudi.config.HoodieWriteConfig
import org.apache.hudi.hive.MultiPartKeysValueExtractor
import org.apache.hudi.hive.HiveSyncConfig
import org.apache.hudi.sync.common.HoodieSyncConfig
```

Escribir en un conjunto de datos de Hudi

En el siguiente ejemplo, se muestra cómo crear DataFrame y escribir como conjunto de datos de Hudi.

> **Note**
>
> Para pegar muestras de código en el shell de Spark, escriba **`:paste`** en el símbolo del sistema, pegue el ejemplo y, a continuación, pulse **`CTRL`** + **`D`**.

Cada vez que escribas DataFrame a en un conjunto de datos de Hudi, debes especificarlo. `DataSourceWriteOptions` Es probable que muchas de estas opciones sean idénticas en varias operaciones de escritura. En el ejemplo siguiente se especifican las opciones comunes mediante la variable *`hudiOptions`*, que se usa en los ejemplos posteriores.

Escribir con Scala con Amazon EMR 6.7 y versiones posteriores

> **Note**
>
> Amazon EMR 6.7.0 usa Apache Hudi 0.11.0-amzn-0, que contiene mejoras significativas con respecto a las versiones anteriores de Hudi. Para obtener más información, consulte la Guía de migración a Apache Hudi 0.11.0. Los ejemplos de esta pestaña reflejan estos cambios.

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TBL_NAME.key -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE.key -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
```

```
  HoodieSyncConfig.META_SYNC_PARTITION_EXTRACTOR_CLASS.key ->
 "org.apache.hudi.hive.MultiPartKeysValueExtractor",
  HoodieSyncConfig.META_SYNC_ENABLED.key -> "true",
  HiveSyncConfig.HIVE_SYNC_MODE.key -> "hms",
  HoodieSyncConfig.META_SYNC_TABLE_NAME.key -> "tableName",
  HoodieSyncConfig.META_SYNC_PARTITION_FIELDS.key -> "creation_date"
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,"insert")
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Escribir con Scala con Amazon EMR 6.6 y versiones anteriores

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TABLE_NAME -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE_OPT_KEY -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY ->
 classOf[MultiPartKeysValueExtractor].getName
)

// Write the DataFrame as a Hudi dataset
```

```
(inputDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.INSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Escribe usando PySpark

```
# Create a DataFrame
inputDF = spark.createDataFrame(
    [
        ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
        ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
        ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
        ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
        ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
        ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z"),
    ],
    ["id", "creation_date", "last_update_time"]
)

# Specify common DataSourceWriteOptions in the single hudiOptions variable
hudiOptions = {
'hoodie.table.name': 'tableName',
'hoodie.datasource.write.recordkey.field': 'id',
'hoodie.datasource.write.partitionpath.field': 'creation_date',
'hoodie.datasource.write.precombine.field': 'last_update_time',
'hoodie.datasource.hive_sync.enable': 'true',
'hoodie.datasource.hive_sync.table': 'tableName',
'hoodie.datasource.hive_sync.partition_fields': 'creation_date',
'hoodie.datasource.hive_sync.partition_extractor_class':
 'org.apache.hudi.hive.MultiPartKeysValueExtractor'
}

# Write a DataFrame as a Hudi dataset
inputDF.write \
.format('org.apache.hudi') \
.option('hoodie.datasource.write.operation', 'insert') \
.options(**hudiOptions) \
.mode('overwrite') \
.save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

> **Note**
>
> Es posible que vea “hoodie” en lugar de Hudi en los ejemplos de código y las notificaciones. El código base de Hudi utiliza ampliamente la antigua palabra “hoodie”.

DataSourceWriteOptions referencia para Hudi

Opción	Descripción
TABLE_NAME	Nombre de la tabla bajo la que se registra el dataset.
TABLE_TYPE_OPT_KEY	Opcional. Especifica si el dataset que se va a crear será de tipo `"COPY_ON_WRITE"` o `"MERGE_ON_READ"`. El valor predeterminado es `"COPY_ON_WRITE"`.
RECORDKEY_FIELD_OPT_KEY	Campo de clave de registro cuyo valor se utilizará como componente `recordKey` de `HoodieKey`. El valor real se obtendrá invocando `.toString()` para el valor del campo. Los campos anidados se pueden especificar mediante la notación de puntos; por ejemplo, `a.b.c`.
PARTITIONPATH_FIELD_OPT_KEY	Campo de la ruta de partición cuyo valor se utilizará como componente `partitionPath` de `HoodieKey`. El valor real se obtendrá invocando `.toString()` para el valor del campo.
PRECOMBINE_FIELD_OPT_KEY	Campo utilizado en la combinación previa antes de la escritura real. Cuando dos registros tienen el mismo valor de clave, Hudi seleccionará el que tiene el valor mayor para el campo de combinación previa, según lo determinado por `Object.compareTo(..)`.

Las siguientes opciones solo son necesarias para registrar la tabla del conjunto de datos de Hudi en su metaalmacén. Si no registra el conjunto de datos de Hudi como tabla en el metaalmacén de Hudi, estas opciones no son necesarias.

DataSourceWriteOptions referencia para Hive

Opción	Descripción
HIVE_DATABASE_OPT_KEY	Base de datos Hive con la que se realizará la sincronización. El valor predeterminado es `"default"` .
HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY	Clase utilizada para extraer valores de los campos de partición a las columnas de partición de Hive.
HIVE_PARTITION_FIELDS_OPT_KEY	Campo del dataset que se utilizará para determinar las columnas de partición de Hive.
HIVE_SYNC_ENABLED_OPT_KEY	Cuando se establece en `"true"`, registra el dataset en el metaalmacén de Apache Hive. El valor predeterminado es `"false"`.
HIVE_TABLE_OPT_KEY	Obligatorio. Nombre de la tabla de Hive con la que se va a realizar la sincronización. Por ejemplo, `"my_hudi_table_cow"` .
HIVE_USER_OPT_KEY	Opcional. Nombre de usuario de Hive que se va a utilizar al sincronizar. Por ejemplo, `"hadoop"`.
HIVE_PASS_OPT_KEY	Opcional. Contraseña de Hive para el usuario especificado por `HIVE_USER_OPT_KEY` .
HIVE_URL_OPT_KEY	Dirección URL del metaalmacén de Hive.

Datos de upsert

En el siguiente ejemplo, se muestra cómo escribir en un para llevar a cabo operaciones upsert. DataFrame A diferencia del ejemplo de inserción anterior, el valor de `OPERATION_OPT_KEY` se establece en `UPSERT_OPERATION_OPT_VAL`. Además, se especifica `.mode(SaveMode.Append)` para indicar que el registro se debe anexar.

Upsert con Scala con Amazon EMR 6.7 y versiones posteriores

> **Note**
>
> Amazon EMR 6.7.0 usa Apache Hudi 0.11.0-amzn-0, que contiene mejoras significativas con respecto a las versiones anteriores de Hudi. Para obtener más información, consulte la Guía de migración a Apache Hudi 0.11.0. Los ejemplos de esta pestaña reflejan estos cambios.

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "upsert")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Upsert con Scala con Amazon EMR 6.6 y versiones posteriores

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Append)
```

```
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Upsert con PySpark

```
from pyspark.sql.functions import lit

# Create a new DataFrame from the first row of inputDF with a different creation_date
 value
updateDF = inputDF.limit(1).withColumn('creation_date', lit('new_value'))

updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

Eliminación de un registro

Para eliminar un registro de forma permanente, puede alterar una carga útil vacía. En este caso, la opción `PAYLOAD_CLASS_OPT_KEY` especifica la clase `EmptyHoodieRecordPayload`. En el ejemplo, se utiliza el mismo DataFrame`updateDF`, que en el ejemplo de upsert para que se especifique el mismo registro.

Eliminación con Scala con Amazon EMR 6.7 y versiones posteriores

> **Note**
>
> Amazon EMR 6.7.0 usa Apache Hudi 0.11.0-amzn-0, que contiene mejoras significativas con respecto a las versiones anteriores de Hudi. Para obtener más información, consulte la Guía de migración a Apache Hudi 0.11.0. Los ejemplos de esta pestaña reflejan estos cambios.

```
(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "delete")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Eliminación con Scala con Amazon EMR 6.6 y versiones anteriores

```
(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .option(DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY,
 "org.apache.hudi.common.model.EmptyHoodieRecordPayload")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Elimine usando PySpark

```
updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .option('hoodie.datasource.write.payload.class',
 'org.apache.hudi.common.model.EmptyHoodieRecordPayload') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

También puede eliminar datos de forma permanente si configura `OPERATION_OPT_KEY` como `DELETE_OPERATION_OPT_VAL` para eliminar todos los registros del conjunto de datos que envíe. Para obtener instrucciones sobre cómo llevar a cabo eliminaciones temporales y para obtener más información sobre cómo eliminar los datos almacenados en las tablas de Hudi, consulte Eliminaciones en la documentación de Apache Hudi.

Leer un conjunto de datos de Hudi

Para recuperar los datos en el momento actual, Hudi lleva a cabo consultas instantáneas de forma predeterminada. El siguiente es un ejemplo de consulta del conjunto de datos escrito en S3 en Escribir en un conjunto de datos de Hudi. Sustituya por *s3://amzn-s3-demo-bucket/myhudidataset* la ruta de la tabla y agregue asteriscos comodín para cada nivel de partición, además de un asterisco adicional. En este ejemplo, hay un nivel de partición, por lo que hemos agregado dos símbolos comodín.

Leer con Scala con Amazon EMR 6.7 y versiones posteriores

> **Note**
>
> Amazon EMR 6.7.0 usa Apache Hudi 0.11.0-amzn-0, que contiene mejoras significativas con respecto a las versiones anteriores de Hudi. Para obtener más información, consulte la Guía de migración a Apache Hudi 0.11.0. Los ejemplos de esta pestaña reflejan estos cambios.

```
val snapshotQueryDF = spark.read
    .format("hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset")
    .show()
```

Leer con Scala con Amazon EMR 6.6 y versiones anteriores

```
(val snapshotQueryDF = spark.read
    .format("org.apache.hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset" + "/*/*"))

snapshotQueryDF.show()
```

Lea usando PySpark

```
snapshotQueryDF = spark.read \
    .format('org.apache.hudi') \
    .load('s3://amzn-s3-demo-bucket/myhudidataset' + '/*/*')

snapshotQueryDF.show()
```

Consultas incrementales

También puede hacer consultas incrementales con Hudi para obtener un flujo de registros que han cambiado desde una fecha de confirmación determinada. Para ello, defina el campo `QUERY_TYPE_OPT_KEY` en `QUERY_TYPE_INCREMENTAL_OPT_VAL`. A continuación, agregue un valor para `BEGIN_INSTANTTIME_OPT_KEY` para obtener todos los registros escritos desde el momento especificado. Las consultas incrementales suelen ser diez veces más eficientes que las consultas por lotes, ya que solo procesan los registros modificados.

Al hacer consultas incrementales, utilice la ruta de la tabla raíz (base) sin los asteriscos comodín que se utilizan en las consultas instantáneas.

> **Note**
>
> Presto no admite consultas incrementales.

Consultas incrementales con Scala

```
(val incQueryDF = spark.read
    .format("org.apache.hudi")
    .option(DataSourceReadOptions.QUERY_TYPE_OPT_KEY,
 DataSourceReadOptions.QUERY_TYPE_INCREMENTAL_OPT_VAL)
    .option(DataSourceReadOptions.BEGIN_INSTANTTIME_OPT_KEY, <beginInstantTime>)
    .load("s3://amzn-s3-demo-bucket/myhudidataset" ))

incQueryDF.show()
```

Consultas incrementales con PySpark

```
readOptions = {
  'hoodie.datasource.query.type': 'incremental',
  'hoodie.datasource.read.begin.instanttime': <beginInstantTime>,
}

incQueryDF = spark.read \
    .format('org.apache.hudi') \
    .options(**readOptions) \
    .load('s3://amzn-s3-demo-bucket/myhudidataset')

incQueryDF.show()
```

Para obtener más información acerca de la lectura de los conjuntos de datos de Hudi, consulte la sección sobre Cómo consultar tablas de Hudi en la documentación de Apache Hudi.

Usar la CLI de Hudi

Puede usar la CLI de Hudi para administrar conjuntos de datos de Hudi y ver información sobre confirmaciones, el sistema de archivos, estadísticas y otros datos. También puede utilizar la CLI para

realizar compactaciones manualmente, programarlas o cancelarlas otras programadas. Para obtener más información, consulte Interacción mediante la CLI en la documentación de Apache Hudi.

Para iniciar la CLI de Hudi y conectarse a un conjunto de datos

1. Conéctese al nodo principal utilizando SSH. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.
2. En la línea de comandos, escriba `/usr/lib/hudi/cli/bin/hudi-cli.sh`.

   El símbolo del sistema cambia a `hudi->`
3. Para conectarse al conjunto de datos, utilice el siguiente comando. Reemplace *`s3://amzn-s3-demo-bucket/myhudidataset`* por la ruta del conjunto de datos con el que desea trabajar. El valor que usamos es el mismo que el valor establecido en ejemplos anteriores.

   ```
   connect --path s3://amzn-s3-demo-bucket/myhudidataset
   ```

   El símbolo del sistema cambia para incluir el dataset al que está conectado, como se muestra en el siguiente ejemplo.

   ```
   hudi:myhudidataset->
   ```

De forma predeterminada, el `hudi-cli.sh` script de las versiones 7.3.0 a 7.8.0 de Amazon EMR utiliza. `hudi-cli-bundle.jar` Si tiene dificultades, puede volver a la CLI Hudi clásica con el siguiente comando:

```
/usr/lib/hudi/cli/bin/hudi-cli.sh --cliBundle false
```

Este comando ejecuta el script `hudi-cli.sh`, establece el indicador `--cliBundle` e indica a la CLI que utilice los archivos JAR individuales en lugar del JAR agrupado. De forma predeterminada, `--cliBundle` se establece en true, lo que significa que la CLI utiliza el JAR agrupado en su lugar.

Uso de Amazon EMR 7.9.0 y versiones posteriores

> **Note**
>
> El script hudi-cli.sh quedó obsoleto en la versión 7.9.0 y posteriores de EMR. La versión 7.9.0 y superior de Amazon EMR usa .jar. hudi-cli-bundle

Para iniciar la CLI de Hudi y conectarse a un conjunto de datos:

1. Conéctese al nodo principal utilizando SSH. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.
2. En la línea de comandos, escriba/usr/lib/hudi/cli-bundle/bin/hudi- cli-with-bundle .sh o simplemente escriba hudi-cli-with-bundleo >hudi-cli.

   La línea de comandos cambia a hudi- >.
3. Para conectarse al conjunto de datos, utilice el siguiente comando. Reemplace s3://amzn-s3-demo-bucket/myhudidataset por la ruta del conjunto de datos con el que desea trabajar. El valor que usamos es el mismo que el valor establecido en ejemplos anteriores.

   ```
   connect --path s3://amzn-s3-demo-bucket/myhudidataset
   ```
4. El símbolo del sistema cambia para incluir el dataset al que está conectado, como se muestra en el siguiente ejemplo.

   ```
   hudi:myhudidataset->
   ```

Historial de versiones de Hudi

En la siguiente tabla, se muestra la versión de Hudi incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Hudi

Etiqueta de versión de Amazon EMR	Versión de Hudi	Componentes instalados con Hudi
emr-7.9.0	0.15.0-amzn-6	Not available.
emr-7.8.0	0.15.0-amzn-5	Not available.
emr-7.7.0	0.15.0-amzn-4	Not available.

Etiqueta de versión de Amazon EMR	Versión de Hudi	Componentes instalados con Hudi
emr-7.6.0	0.15.0-amzn-3	Not available.
emr-7.5.0	0.15.0-amzn-2	Not available.
emr-7.4.0	0.15.0-amzn-1	Not available.
emr-7.3.0	0.15.0-amzn-0	Not available.
emr-7.2.0	0.14.1-amzn-1	Not available.
emr-5.36.2	0.10.1-amzn-1	Not available.
emr-7.1.0	0.14.1-amzn-0	Not available.
emr-7.0.0	0.14.0-amzn-1	Not available.
emr-6.15.0	0.14.0-amzn-0	Not available.
emr-6.14.0	0.13.1-amzn-2	Not available.
emr-6.13.0	0.13.1-amzn-1	Not available.
emr-6.12.0	0.13.1-amzn-0	Not available.
emr-6.11.1	0.13.0-amzn-0	Not available.
emr-6.11.0	0.13.0-amzn-0	Not available.
emr-6.10.1	0.12.2-amzn-0	Not available.
emr-6.10.0	0.12.2-amzn-0	Not available.
emr-6.9.1	0.12.1-amzn-0	Not available.
emr-6.9.0	0.12.1-amzn-0	Not available.
emr-6.8.1	0.11.1-amzn-0	Not available.
emr-6.8.0	0.11.1-amzn-0	Not available.

Etiqueta de versión de Amazon EMR	Versión de Hudi	Componentes instalados con Hudi
emr-6.7.0	0.11.0-amzn-0	Not available.
emr-5.36.1	0.10.1-amzn-1	Not available.
emr-5.36.0	0.10.1-amzn-1	Not available.
emr-6.6.0	0.10.1-amzn-0	Not available.
emr-5.35.0	0.9.0-amzn-2	Not available.
emr-6.5.0	0.9.0-amzn-1	Not available.
emr-6.4.0	0.8.0-amzn-0	Not available.
emr-6.3.1	0.7.0-amzn-0	Not available.
emr-6.3.0	0.7.0-amzn-0	Not available.
emr-6.2.1	0.6.0-amzn-1	Not available.
emr-6.2.0	0.6.0-amzn-1	Not available.
emr-6.1.1	0.5.2-incubating-amzn-2	Not available.
emr-6.1.0	0.5.2-incubating-amzn-2	Not available.
emr-6.0.1	0.5.0-incubating-amzn-1	Not available.
emr-6.0.0	0.5.0-incubating-amzn-1	Not available.
emr-5.34.0	0.9.0-amzn-0	Not available.
emr-5.33.1	0.7.0-amzn-1	Not available.
emr-5.33.0	0.7.0-amzn-1	Not available.
emr-5.32.1	0.6.0-amzn-0	Not available.
emr-5.32.0	0.6.0-amzn-0	Not available.

Etiqueta de versión de Amazon EMR	Versión de Hudi	Componentes instalados con Hudi
emr-5.31.1	0.6.0-amzn-0	Not available.
emr-5.31.0	0.6.0-amzn-0	Not available.
emr-5.30.2	0.5.2-incubating	Not available.
emr-5.30.1	0.5.2-incubating	Not available.
emr-5.30.0	0.5.2-incubating	Not available.
emr-5.29.0	0.5.0-incubating	Not available.
emr-5.28.1	0.5.0-incubating	Not available.
emr-5.28.0	0.5.0-incubating	Not available.

Hue

Hue (Hadoop User Experience) es una interfaz gráfica de usuario de código abierto basada en web para su uso con Amazon EMR y Apache Hadoop. Hue agrupa diferentes proyectos del ecosistema de Hadoop en una interfaz configurable. Amazon EMR también ha agregado las personalizaciones específicas de Hue en Amazon EMR. Hue actúa como front-end para las aplicaciones que se ejecutan en el clúster, lo que le permite interactuar con las aplicaciones mediante una interfaz que ya conozca o que sea más fácil de utilizar. Las aplicaciones en Hue, como los editores de Hive y Pig, sustituyen la necesidad de iniciar sesión en el clúster para ejecutar scripts de forma interactiva usando el shell correspondiente de cada aplicación. Una vez lanzado un clúster, puede interactuar completamente con las aplicaciones mediante Hue o una interfaz similar. Para obtener información adicional sobre Hue, consulte http://gethue.com.

Hue se instala de forma predeterminada cuando lanza el clúster mediante la consola de Amazon EMR. Puede optar por no instalar Hue mediante Opciones avanzadas en la consola de Amazon EMR cuando lance un clúster o si especifica explícitamente la opción `--applications` y, al omitir Hue, cuando utilice `create-cluster` desde la AWS CLI.

Temas

- Información de la versión de Hue
- Características de Hue admitidas y no admitidas en Amazon EMR
- Consideraciones
- Conexión a la interfaz de usuario web de Hue
- Uso de Hue con una base de datos remota en Amazon RDS
- Configuraciones avanzadas para Hue
- Historial de versiones de Hue

Información de la versión de Hue

Versión 7.9.0

En la siguiente tabla, se muestra la versión de Hue incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Hue.

Para ver la versión de los componentes instalados con Hue en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Hue para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-7.9.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Versión de Hue para 6.15.0

En la tabla siguiente, se muestra la versión de Hue incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Hue.

Para ver la versión de los componentes instalados con Hue en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Hue para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.15.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server,

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
		mariadb-server, oozie-client, oozie-server

Versión de Hue para 5.36.2

En la tabla siguiente, se muestra la versión de Hue incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Hue.

Para ver la versión de los componentes instalados con Hue en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Hue para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.36.2	Hue 4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Características de Hue admitidas y no admitidas en Amazon EMR

Con Amazon EMR 7.0.0 y versiones posteriores, Hue requiere la versión de Python 3.9 o posterior. Cuando utilice el cifrado en tránsito, deberá utilizar una clave de cifrado de 2048 bits.

- Amazon S3 y Hadoop File System (HDFS) Browser

  - Con los permisos adecuados, puede examinar y mover datos entre el almacenamiento efímero de HDFS y los buckets de S3 que pertenecen a su cuenta.
  - Los superusuarios de Hue pueden acceder, de forma predeterminada, a todos los archivos para los que tienen autorización de acceso los roles de IAM de Amazon EMR. Los usuarios recién creados no obtienen automáticamente permisos de acceso al navegador de archivos de Amazon S3, por lo que es necesario activar los permisos `filebrowser.s3_access` para su grupo.
- Hive: ejecute consultas interactivas en los datos. También es una forma útil de crear prototipos mediante programación o en lotes de consultas.
- Pig: ejecute scripts en los datos o emita comandos interactivos.
- Oozie: cree y supervise flujos de trabajo de Oozie.
- Metastore Manager: vea y manipule el contenido del almacén de metadatos de Hive (importar, crear, soltar, etc.).
- Navegador de trabajos: vea el estado de los trabajos de Hadoop enviados.
- Administración de usuarios: administre cuentas de Hue para integrar a los usuarios de LDAP con Hue.
- AWS Ejemplos de: existen varios «ready-to-run» que procesan datos de muestra de varios AWS servicios de con aplicaciones en Hue. Al iniciar sesión en Hue, accede a una aplicación de inicio de Hue donde los ejemplos están preinstalados.
- Livy Server solo se admite en la versión 5.9.0 y posteriores de Amazon EMR.
- Para utilizar Hue Notebook para Spark, debe instalar Hue con Livy y Spark.
- Trino/Presto: Hue admite la conexión a los servidores de datos Trino y Presto. Con Amazon EMR 7.0.0 y versiones posteriores, esta conectividad se facilita con `trino-python-client` y `presto-python-client` conectores. Las versiones anteriores a la 7.0.0 utilizan un conector de JDBC a través de una puerta de enlace de JDBC.
- El panel de Hue no se admite.
- No se admite PostgreSQL.

Consideraciones

Tenga en cuenta las siguientes limitaciones al ejecutar Hue en Amazon EMR.

Rendimiento con tablas de metadatos de Hue de gran tamaño

Si la base de datos de metadatos de Hue se hace demasiado grande, es posible que el rendimiento disminuya. Para comprobar el tamaño de las tablas, primero conéctese al nodo principal del EC2 clúster de Amazon EMR en Amazon con SSH y ejecute el comando `sudo mysql -u root` para iniciar la CLI de MySQL. Para obtener el tamaño de la tabla, ejecute la consulta `SELECT COUNT(*) FROM hue.<table_name>`. Consulte lo siguiente para ver lo que *`<table_name>`* puede ser:

- desktop_document
- desktop_document2
- oozie_job
- beeswax_savedquery
- beeswax_session
- beeswax_queryhistory

Si al ejecutar esa consulta se obtiene un recuento superior a 100 000, debe ejecutar el siguiente comando de limpieza para eliminar los registros antiguos.

```
cd /opt/cloudera/parcels/CDH/lib/hue # Hue home directory ./build/env/bin/hue
 desktop_document_cleanup
```

Para obtener más información sobre cómo limpiar la base de datos, consulte la página de referencia de la documentación de Hue.

Hue no limpia automáticamente las tablas, pero las versiones 5.12.0 y posteriores de Amazon EMR proporcionan un método para eliminar documentos antiguos de las tablas. Cree el siguiente script de shell y ejecútelo como un paso en un clúster de Amazon EMR con un parámetro entero que represente el número máximo de días durante los que se deben conservar los documentos en la base de datos de metadatos.

```
#!/bin/bash
if grep isMaster /mnt/var/lib/info/instance.json | grep false;
then
  echo "This is not the primary node; do nothing, exiting"
  exit 0
fi
while [ ! -f /usr/lib/hue/desktop/core/src/desktop/management/commands/
desktop_document_cleanup.py ]
```

```
do
  sleep 1
done
sudo systemctl stop hue.service
sudo sed -i 's+  LOG.warn+  # LOG.warn+g' /usr/lib/hue/desktop/core/src/desktop/
management/commands/desktop_document_cleanup.py
sudo /usr/lib/hue/build/env/bin/hue desktop_document_cleanup --keep-days $1
sudo systemctl start hue.service
```

Incompatibilidad entre versiones de Hue

Si utiliza la misma base de datos de metadatos de Hue en varios clústeres habilitados para Hue, le recomendamos que estos clústeres ejecuten la misma versión de Hue. Las distintas versiones de Hue pueden tener diferentes esquemas para la base de datos de metadatos de Hue. El uso de la misma base de datos para diferentes versiones puede provocar un error en la instalación de Hue. Por ejemplo, intentar usar la misma base de datos para dos clústeres con las versiones 4.10.0 y 4.11.0 instaladas puede provocar errores de inicio de sesión en los usuarios que intenten iniciar sesión en el clúster 4.10.0.

Para obtener una lista de las etiquetas de lanzamiento de Amazon EMR y las correspondientes versiones instaladas de Hue, consulte el historial de versiones de Hue.

Conexión a la interfaz de usuario web de Hue

La conexión con la interfaz de usuario web de Hue es equivalente a conectar a cualquier interfaz HTTP alojada en el nodo maestro de un clúster. El siguiente procedimiento describe cómo acceder a la interfaz de usuario de Hue. Para obtener más información, consulte Ver interfaces web alojadas en clústeres de EMR en la Guía de administración de Amazon EMR.

Para ver la interfaz de usuario de Hue

1. Siga las instrucciones para configurar un túnel SSH hacia el nodo maestro mediante el reenvío dinámico de puertos en la Guía de administración de Amazon EMR.
2. Escriba la siguiente dirección en el navegador para abrir la interfaz web de Hue: por ejemplo, `http://`*`master public DNS`*`:8888` dónde *`master public dns`* está el nombre de DNS público del nodo principal del clúster`ec2-11-22-333-44.compute-1.amazonaws.com`.
3. En la pantalla de inicio de sesión de Hue, si es el administrador que inicia sesión por primera vez, escriba un nombre de usuario y una contraseña para crear su supercuenta de Hue y, a

continuación, seleccione Crear cuenta. En caso contrario, escriba su nombre de usuario y la contraseña y seleccione Create account (Crear cuenta) o introduzca las credenciales que le proporcionó el administrador.

Uso de Hue con una base de datos remota en Amazon RDS

De forma predeterminada, la información de usuario de Hue y las historias de consulta se almacenan en una base de datos MySQL local en el nodo principal. También puede crear uno o varios clústeres habilitados para Hue con una configuración almacenada en Amazon S3 y una base de datos MySQL en Amazon Relational Database Service (Amazon RDS). Esto le permite conservar información de usuario y el historial de consultas creado por Hue sin mantener el clúster de Amazon EMR en ejecución. Le recomendamos utilizar el cifrado del servidor de Amazon S3 para almacenar el archivo de configuración.

En primer lugar, cree la base de datos remota para Hue.

Para crear una base de datos MySQL externa

1. Abra la consola de Amazon RDS en https://console.aws.amazon.com/rds/.
2. Haga clic en Launch a DB Instance (Lanzar una instancia de base de datos).
3. Elija MySQL y haga clic en Select (Seleccionar).
4. Deje la selección predeterminada de Multi-AZ Deployment and Provisioned IOPS Storage (Implementación Multi-AZ y almacenamiento de IOPS provisionadas) y haga clic en Next (Siguiente).
5. Deje los valores predeterminados de Instance Specifications (Especificaciones de la instancia), especifique Settings (Configuración) y haga clic en Next (Siguiente).
6. En la página Configure Advanced Settings, elija un grupo de seguridad y un nombre de base de datos adecuados. El grupo de seguridad que utilice debe, al menos, permitir el acceso TCP de entrada al puerto 3306 desde el nodo principal del clúster. Si no ha creado su clúster en este punto, puede permitir que todos los hosts se conecten al puerto 3306 y ajustar el grupo de seguridad después de haber lanzado el clúster. Haga clic en Launch DB Instance (Lanzar instancia de base de datos).
7. En el panel de RDS, seleccione Instances (Instancias) y seleccione la instancia que acaba de crear. Cuando la base de datos esté disponible, anote el dbname, nombre de usuario, contraseña y nombre de host de la instancia de RDS. Utilizará esta información al crear y configurar el clúster.

Para especificar una base de datos MySQL externa para Hue al lanzar un clúster mediante el AWS CLI

Para especificar una base de datos MySQL externa para Hue al lanzar un clúster mediante AWS CLI, utilice la información que anotó al crear la instancia de RDS para configurarla `hue.ini` con un objeto de configuración.

> **Note**
>
> Puede crear varios clústeres que utilicen la misma base de datos externa, pero cada clúster compartirá el historial de consultas y la información de usuario.

- Utilice el AWS CLI, cree un clúster con Hue instalado, utilizando la base de datos externa que creó y haciendo referencia a un archivo de configuración con una clasificación de configuración para Hue que especifique las propiedades de la base de datos. El siguiente ejemplo crea un clúster con Hue instalado, que hace referencia a un archivo de configuración de Amazon S3, `myConfig.json`, que especifica la configuración de la base de datos.

  > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
   Name=Hive \
  --instance-type m5.xlarge --instance-count 3 \
  --configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
  myConfig.json --use-default-roles
  ```

  A continuación, se muestra contenido de ejemplo del archivo `myConfig.json`. Sustituya *dbnameusername*,*password*, y por *RDS instance hostname* los valores que anotó anteriormente en el panel de control de RDS.

  ```
  [{
    "Classification": "hue-ini",
    "Properties": {},
  ```

```
  "Configurations": [
    {
      "Classification": "desktop",
      "Properties": {},
      "Configurations": [
        {
          "Classification": "database",
          "Properties": {
            "name": "dbname",
            "user": "username",
            "password": "password",
            "host": "RDS instance hostname",
            "port": "3306",
            "engine": "mysql"
          },
          "Configurations": []
        }
      ]
    }
  ]
}]
```

Solución de problemas

En el caso de conmutación por error de Amazon RDS

Es posible que los usuarios puedan encontrar retrasos a la hora de ejecutar una consulta, ya que la instancia de base de datos de Hue no responde o está en proceso de conmutación por error. A continuación, especificamos algunos hechos y directrices importantes para este problema:

- Si inicia sesión en la consola de Amazon RDS, puede buscar eventos de conmutación por error. Por ejemplo, para ver si una conmutación por error está en curso o se ha producido busque eventos como "Multi-AZ instance failover started" y "Multi-AZ instance failover completed".
- Se tarda unos 30 segundos en que una instancia de RDS complete una conmutación por error.
- Si está viendo longer-than-normal respuestas a las consultas en Hue, intente volver a ejecutar la consulta.

Configuraciones avanzadas para Hue

Esta sección incluye los siguientes temas.

Temas

- Configurar Hue para usuarios de LDAP

Configurar Hue para usuarios de LDAP

La integración con LDAP permite a los usuarios iniciar sesión en Hue utilizando credenciales existentes almacenadas en un directorio LDAP. Cuando se integra Hue con LDAP, no es necesario administrar de forma independiente la información de usuario en Hue. La información que aparece a continuación muestra la integración de Hue con Microsoft Active Directory, pero las opciones de configuración son análogas para cualquier directorio LDAP.

La autenticación LDAP primero se enlaza al servidor y establece la conexión. A continuación, la conexión establecida se utiliza para las consultas posteriores para buscar información de usuario de LDAP. A menos que el servidor de Active Directory permita conexiones anónimas, es necesario establecer una conexión con un nombre distinguido y contraseña de enlace. El nombre distinguido (o DN) de enlace se define mediante el ajuste de configuración `bind_dn`. La contraseña de enlace se define mediante el ajuste de configuración `bind_password`. Hue tiene dos formas de enlazar las solicitudes de LDAP: enlace de búsqueda y enlace directo. El método preferido para utilizar Hue con Amazon EMR es el enlace de búsqueda.

Cuando se utiliza el enlace de búsqueda con Active Directory, Hue utiliza el atributo de nombre de usuario (definido por `user_name_attr config`) para buscar el atributo que hay que recuperar desde el nombre distinguido (o DN) base. El enlace de búsqueda resulta útil cuando el usuario Hue no conoce el DN completo.

Por ejemplo, es posible que tenga `user_name_attr config` definido para utilizar el nombre común (o CN). En ese caso, el servidor de Active Directory utiliza el nombre de usuario de Hue proporcionado durante el inicio de sesión para buscar en el árbol de directorios un nombre común que coincida, empezando por el nombre distinguido base. Si se encuentra el nombre común para el usuario de Hue, el servidor devuelve el nombre distinguido del usuario. A continuación, Hue construye un nombre distinguido utilizado para autenticar al usuario realizando una operación de enlace.

> **Note**
>
> El enlace de búsqueda busca los nombres de usuario en todos los subárboles de directorio empezando en el nombre distinguido base. El nombre distinguido base especificado en la configuración de LDAP de Hue debe ser el elemento primario más cercano del nombre de usuario o el rendimiento de autenticación de LDAP podría verse afectado.

Cuando se utiliza el enlace directo con Active Directory, se debe utilizar el `nt_domain` o `ldap_username_pattern` exacto para la autenticación. Cuando se utiliza el enlace directo, si se define atributo de dominio nt (definido por el ajuste de configuración `nt_domain`), se crea una plantilla de nombre distinguido de usuario utilizando el formulario: `<login username>@nt_domain`. Esta plantilla se utiliza para buscar en todos los subárboles de directorio empezando en el nombre distinguido base. Si el dominio nt no está configurado, Hue busca un patrón de nombre distinguido exacto para el usuario (definido por el ajuste de configuración `ldap_username_pattern`). En esta instancia, el servidor busca un valor `ldap_username_pattern` coincidente en todos los subárboles de directorios empezando por el nombre distinguido base.

Para lanzar un clúster con propiedades de LDAP para Hue con la AWS CLI

- Para especificar las propiedades de LDAP para `hue-ini`, cree un clúster con Hue instalado y haga referencia a un archivo json con propiedades de configuración de LDAP. A continuación se muestra un comando de ejemplo, que hace referencia a un archivo de configuración `myConfig.json` almacenado en Amazon S3.

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
   Name=Hive \
  --instance-type m5.xlarge --instance-count 3 --configurations https://
  s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json.
  ```

  A continuación, se muestra contenido de ejemplo de `myConfig.json`.

  ```
  [
      {
          "Classification": "hue-ini",
          "Properties": {},
          "Configurations": [
              {
  ```

```
                "Classification": "desktop",
                "Properties": {},
                "Configurations": [
                    {
                        "Classification": "ldap",
                        "Properties": {},
                        "Configurations": [
                            {
                                "Classification": "ldap_servers",
                                "Properties": {},
                                "Configurations": [
                                    {
                                        "Classification": "yourcompany",
                                        "Properties": {
                                            "base_dn":
 "DC=yourcompany,DC=hue,DC=com",
                                            "ldap_url": "ldap://ldapurl",
                                            "search_bind_authentication": "true",
                                            "bind_dn":
 "CN=hue,CN=users,DC=yourcompany,DC=hue,DC=com",
                                            "bind_password": "password"
                                        },
                                        "Configurations": []
                                    }
                                ]
                            }
                        ]
                    },
                    {
                        "Classification": "auth",
                        "Properties": {
                         "backend": "desktop.auth.backend.LdapBackend"
                              }
                    }
                ]
            }

        ]
    }
]
```

> **Note**
>
> Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Para ver la configuración de LDAP en Hue

1. Compruebe que tenga una conexión de VPN activa o túnel SSH en el nodo maestro del clúster de Amazon EMR. A continuación, en su navegador, escriba:8888 *master-public-dns* para abrir la interfaz web de Hue.
2. Inicie sesión con sus credenciales de administrador de Hue. Si se abre la ventana Did you know? (¿Sabía que...?) haga clic en Got it, prof! (¡Ya lo sé!) para cerrarla.
3. Haga clic en el icono Hue en la barra de herramientas.
4. En la página Acerca de Hue, seleccione Configuración.
5. En la sección Configuration Sections and Variables (Secciones de configuración y variables), haga clic en Desktop (Escritorio).
6. Desplácese a la sección ldap para ver su configuración.

Historial de versiones de Hue

En la siguiente tabla, se muestra la versión de Hue incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Hue

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-7.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-cl

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
		ient, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-7.6.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.5.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-7.4.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.3.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-7.2.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.2	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-7.1.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.0.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.15.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.14.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.13.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.12.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.10.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.10.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.9.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.9.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.8.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.8.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.7.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.36.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.6.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.35.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.3.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.3.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.2.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.2.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.1.1	4,71	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.1.0	4,71	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.0.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-6.0.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.34.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.33.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.33.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.32.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.32.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.31.1	4,71	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.31.0	4,71	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.30.2	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.30.1	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.30.0	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.29.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.28.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.28.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.27.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.27.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.26.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.25.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.24.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.24.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.23.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.23.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.22.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.21.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.20.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.20.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.19.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.19.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.18.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.18.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.17.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.17.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.17.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.16.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.16.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.15.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.15.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.14.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.13.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.13.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.12.3	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.12.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.11.4	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.3	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.11.2	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.11.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.10.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.10.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.9.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.9.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.8.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.8.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.7.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.7.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.6.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.6.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.4	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.5.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.4.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.4.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.2	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.3.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.2.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.1.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.1.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.0.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-5.0.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.9.6	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.9.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.8.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.8.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.7.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.7.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.7.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.6.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.6.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.5.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.4.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Etiqueta de versión de Amazon EMR	Versión de Hue	Componentes instalados con Hue
emr-4.3.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.2.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.1.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Iceberg

Apache Iceberg es un formato de tabla abierta para grandes conjuntos de datos de Amazon Simple Storage Service (Amazon S3). Ofrece un rendimiento de consultas rápido en tablas grandes, confirmaciones atómicas, escrituras simultáneas y una evolución de tablas compatible con SQL. A partir de Amazon EMR 6.5.0, puede usar Apache Spark 3 en clústeres de Amazon EMR con el formato de tabla de Iceberg.

En la siguiente tabla, se muestra la versión de Iceberg incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Iceberg.

Para ver la versión de los componentes instalados con Iceberg en esta versión, consulte Versiones de componentes de la versión 7.9.0.

Información sobre la versión de Iceberg para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Iceberg	Componentes instalados con Iceberg
emr-7.9.0	Iceberg 1.7.1-amzn-2	Not available.

En la tabla siguiente, se muestra la versión de Iceberg incluida en la última versión de la serie 6.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Iceberg.

Para ver la versión de los componentes instalados con Iceberg en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Iceberg para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Iceberg	Componentes instalados con Iceberg
emr-6.15.0	Iceberg 1.4.0-amzn-0	Not available.

Temas

- Cómo funciona Iceberg
- Uso de un clúster con Iceberg

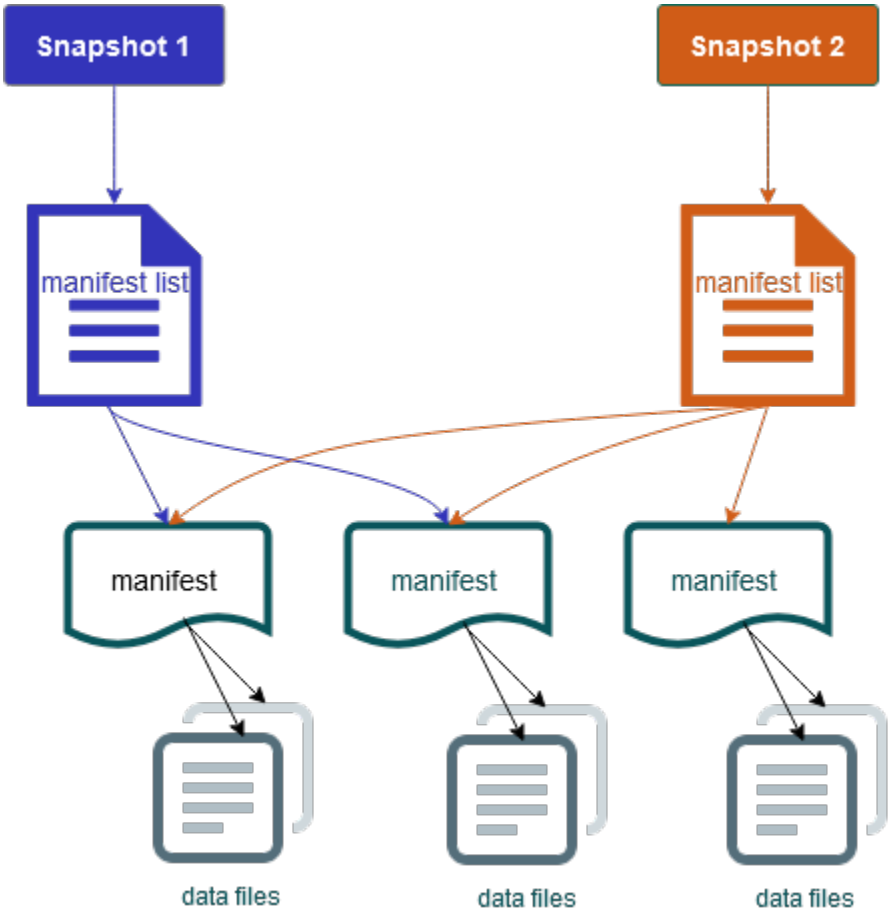
- Historial de versiones de Iceberg

Cómo funciona Iceberg

Iceberg rastrea los archivos de datos individuales en una tabla en lugar de directorios. De esta forma, los escritores pueden crear archivos de datos in situ (los archivos no se mueven ni se modifican). Además, los escritores solo pueden agregar archivos a la tabla mediante una confirmación explícita. El estado de la tabla se mantiene en los archivos de metadatos. Todos los cambios en el estado de la tabla crean un nuevo archivo de metadatos que reemplaza atómicamente a los metadatos anteriores. El archivo de metadatos de la tabla rastrea el esquema de la tabla, la configuración de las particiones y otras propiedades.

También incluye instantáneas del contenido de la tabla. Cada instantánea es un conjunto completo de archivos de datos de la tabla en un momento dado. Las instantáneas se muestran en el archivo de metadatos, pero los archivos de una instantánea se almacenan en archivos de manifiesto independientes. Las transiciones atómicas de un archivo de metadatos de una tabla al siguiente aíslan las instantáneas. Los lectores utilizan la instantánea que estaba actualizada cuando cargaron los metadatos de la tabla. Los lectores no se ven afectados por los cambios hasta que actualizan y eligen una nueva ubicación de metadatos. Los archivos de datos de las instantáneas se almacenan en uno o varios archivos de manifiesto que contienen una fila para cada archivo de datos de la tabla, sus datos de partición y sus métricas. Una instantánea es la unión de todos los archivos de sus manifiestos. Los archivos de manifiesto también se pueden compartir entre instantáneas para evitar reescribir los metadatos que cambian con poca frecuencia.

Diagrama de instantáneas de Iceberg



Iceberg ofrece las siguientes características:

- Admite transacciones ACID y viajes en el tiempo en su lago de datos de Amazon S3.
- Los reintentos confirmados se benefician de las ventajas de rendimiento de una simultaneidad optimista.
- La resolución de conflictos a nivel de archivo da como resultado una alta simultaneidad.
- Con estadísticas mínimas y máximas por columna en los metadatos, puede omitir archivos, lo que mejora el rendimiento de las consultas selectivas.
- Puede organizar las tablas en diseños de particiones flexibles, y la evolución de las particiones permite actualizar los esquemas de particiones. De este modo, las consultas y los volúmenes de datos pueden cambiar sin tener que depender de directorios físicos.
- Admite evolución de los esquemas y cumplimiento.
- Las tablas de Iceberg actúan como receptores idempotentes y orígenes reproducibles. Esto permite la transmisión y la compatibilidad de lotes con canalizaciones que se ejecutan exactamente una sola vez. Los receptores idempotentes rastrean las operaciones de escritura que

han tenido éxito en el pasado. Por lo tanto, el receptor puede volver a solicitar datos en caso de error y descartarlos si se han enviado varias veces.

- Consulte el historial y el linaje, incluida la evolución de la tabla, el historial de operaciones y las estadísticas de cada confirmación.
- Migre desde un conjunto de datos existente con una selección de formatos de datos (Parquet, ORC, Avro) y motor de análisis (Spark, Trino, PrestoDB, Flink, Hive).

Uso de un clúster con Iceberg

Esta sección incluye información sobre cómo usar Iceberg con Spark, Trino, Flink y Hive.

Uso de un clúster de Iceberg con Spark

A partir de la versión 6.5.0 de Amazon EMR, puede usar Iceberg con su clúster de Spark sin necesidad de incluir acciones de arranque. Para las versiones 6.4.0 y anteriores de Amazon EMR, puede utilizar una acción de arranque para preinstalar todas las dependencias necesarias.

En este tutorial, utilizará el AWS CLI para trabajar con Iceberg en un clúster de Amazon EMR Spark. Para usar la consola para crear un clúster con Iceberg instalado, siga los pasos que se indican en Crear un lago de datos de Apache Iceberg con Amazon Athena, Amazon EMR y AWS Glue.

Crear un clúster de Iceberg

Puede crear un clúster con Iceberg instalado mediante la AWS Management Console, la AWS CLI o la API de Amazon EMR. En este tutorial, utilizará el AWS CLI para trabajar con Iceberg en un clúster de Amazon EMR. Para usar la consola para crear un clúster con Iceberg instalado, siga los pasos que se indican en Crear un lago de datos de Apache Iceberg con Amazon Athena, Amazon EMR y AWS Glue.

Para usar Iceberg en Amazon EMR con AWS CLI el, cree primero un clúster con los siguientes pasos. Para obtener información sobre cómo especificar la clasificación de Iceberg mediante la AWS CLI, consulte o. Proporcione una configuración mediante AWS CLI al crear un clúster Proporcione una configuración mediante el SDK de Java al crear un clúster

1. Cree un archivo `configurations.json`, con el siguiente contenido:

```
[{
    "Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
```

```
}]
```

2. A continuación, cree un clúster con la siguiente configuración. Sustituya la ruta del bucket de Amazon S3 de ejemplo y el ID de subred por los suyos propios.

```
aws emr create-cluster --release-label emr-6.5.0 \
--applications Name=Spark \
--configurations file://configurations.json \
--region us-east-1 \
--name My_Spark_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Como alternativa, puede crear un clúster de Amazon EMR que incluya la aplicación de Spark e incluir el archivo `/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` como una dependencia JAR en un trabajo de Spark. Para obtener más información, consulte Envío de aplicaciones.

Para incluir el jar como una dependencia en un trabajo de Spark, agregue la siguiente propiedad de configuración a la aplicación de Spark:

```
--conf "spark.jars=/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar"
```

Para obtener más información sobre las dependencias de los trabajos de Spark, consulte Administración de dependencias en el documento de Apache Spark Ejecución de Spark en Kubernetes.

Inicio de una sesión de Spark para Iceberg

En los siguientes ejemplos, se muestra cómo iniciar el intérprete de comandos interactivo de Spark, utilizar el envío de Spark o usar Cuadernos de Amazon EMR para trabajar con Amazon EMR.

spark-shell

1. Conéctese al nodo principal utilizando SSH. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.

2. Introduzca el siguiente comando para iniciar el shell de Spark. Para usar la PySpark carcasa, sustitúyala por`spark-shell`. `pyspark`

```
spark-shell \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. Conéctese al nodo principal utilizando SSH. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.
2. Ingrese el siguiente comando para iniciar la sesión de Spark para Iceberg.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

Para iniciar una sesión de Spark con los Cuadernos de Amazon EMR, configure su sesión de Spark con el comando mágico `%%configure` de su cuaderno de Amazon EMR, como en el siguiente ejemplo. Para obtener más información, consulte Usar la magia de Cuadernos de Amazon EMR en la Guía de administración de Amazon EMR.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",
```

```
 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
}
```

CLI

Para inicializar un clúster de Spark mediante la CLI y establecer todas las configuraciones predeterminadas de la sesión de Spark Iceberg, ejecute el siguiente ejemplo. Para obtener más información sobre cómo especificar una clasificación de configuración mediante la API Amazon EMR AWS CLI y Amazon EMR, consulte Configurar aplicaciones.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

Escribir en una tabla de Iceberg

En el siguiente ejemplo, se muestra cómo crear un conjunto de datos Iceberg DataFrame y cómo escribirlo como un conjunto de datos. En los ejemplos, se muestra cómo trabajar con conjuntos de datos mediante el intérprete de comandos de Spark mientras está conectado al nodo maestro usando SSH como usuario predeterminado de Hadoop.

> **Note**
>
> Para pegar muestras de código en el intérprete de comandos de Spark, escriba `:paste` en el símbolo del sistema, pegue el ejemplo y, a continuación, pulse `CTRL+D`.

PySpark

Spark incluye un intérprete de comandos basado en Python, `pyspark`, que puede utilizar para crear prototipos de programas de Spark escritos en Python. Invocar `pyspark` en el nodo maestro.

```
## Create a DataFrame.
data = spark.createDataFrame([
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
],["id", "creation_date", "last_update_time"])

## Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame.
val data = Seq(
("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
).toDF("id", "creation_date", "last_update_time")

// Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")
```

```
data.writeTo("dev.db.iceberg_table").append()
```

Leer desde una tabla de Iceberg

PySpark

```
df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Scala

```
val df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Spark SQL

```
SELECT * from dev.db.iceberg_table LIMIT 10
```

Uso del catálogo de datos de AWS Glue con Spark Iceberg

Puedes conectarte al catálogo de datos de AWS Glue desde Spark Iceberg. En esta sección se muestran diferentes comandos para la conexión.

Conéctate al catálogo AWS Glue predeterminado de tu región predeterminada

En este ejemplo se muestra cómo realizar la conexión mediante el tipo de catálogo Glue. Si no especificas un identificador de catálogo, se utilizará el predeterminado:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Conectarse a un catálogo de AWS Glue con un identificador de catálogo específico

En este ejemplo se muestra cómo conectarse mediante un ID de catálogo:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=AWS Glue catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Este comando se puede utilizar para conectarse a un catálogo de AWS Glue de una cuenta diferente, a un catálogo de RMS o a un catálogo federado.

Uso del catálogo REST (IRC) de Iceberg con Spark Iceberg

En las secciones siguientes se detalla cómo configurar la integración de Iceberg con un catálogo.

Conéctese al AWS terminal IRC de Glue Data Catalog

A continuación se muestra un ejemplo de `spark-submit` comando para usar Iceberg REST:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=glue catalog ID \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=glue endpoint URI/iceberg \
    --conf spark.sql.catalog.my_catalog.rest.sigv4-enabled=true \
    --conf spark.sql.catalog.my_catalog.rest.signing-name=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Para usarlo en un clúster con funciones de ejecución habilitadas, se necesitan los siguientes ajustes de configuración de Spark adicionales:

```
"spark.hadoop.fs.s3.credentialsResolverClass":
 "software.amazon.glue.GlueTableCredentialsResolver",
"spark.hadoop.catalog-impl": "org.apache.iceberg.aws.glue.GlueCatalog",
"spark.hadoop.glue.id": glue catalog ID
"spark.hadoop.glue.endpoint": "glue endpoint"
```

Para ver la lista de URL de terminales de AWS Glue para cada región, consulta Cuotas y puntos de conexión de AWS Glue.

Conéctese a un punto final IRC arbitrario

A continuación se muestra un ejemplo de `spark-submit` comando para usar un punto final de IRC:

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=warehouse name \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=your rest endpoint \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Diferencias de configuración cuando se usa Iceberg SparkCatalog y SparkSessionCatalog

Iceberg ofrece dos formas de crear catálogos de Spark Iceberg. Puedes configurar la configuración de Spark en una o `SparkCatalog` en. `SparkSessionCatalog`

¿Usando Iceberg SparkCatalog

A continuación se muestra el comando que se puede utilizar SparkCatalogcomo catálogo de Spark Iceberg:

```
spark-shell \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog
```

Consideraciones para este enfoque:

- Puede acceder a las tablas de Iceberg, pero no a otras tablas.
- El nombre del catálogo no puede ser spark_catalog. Es el nombre del catálogo inicial de Spark. Siempre se conecta a una metatienda de Hive. Es el catálogo predeterminado en Spark, a menos que el usuario lo sobrescriba mediante. `spark.sql.defaultCatalog`
- Puedes establecer el nombre `spark.sql.defaultCatalog` de tu catálogo para convertirlo en el catálogo predeterminado.

¿Usando Iceberg SparkSessionCatalog

A continuación se muestra el comando que se puede utilizar SparkSessionCatalogcomo catálogo de Spark Iceberg:

```
spark-shell \
    --conf spark.sql.catalog.spark_catalog=org.apache.iceberg.spark.SparkSessionCatalog
 \
    --conf spark.sql.catalog.spark_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.spark_catalog.type=glue
```

Consideraciones para este enfoque:

- Si una tabla no se encuentra como una tabla de Iceberg, Spark intentará comprobar si se trata de una tabla en el metaalmacén de Hive. Consulte Uso del catálogo de datos de AWS Glue como catálogo de Hive para obtener más información.
- El nombre del catálogo debe ser spark_catalog.

Uso de extensiones Iceberg Spark

Iceberg ofrece la extensión Spark `org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions` que los usuarios pueden configurar a través de la configuración de extensiones de Spark. `spark.sql.extensions` Las extensiones permiten funciones clave de Iceberg, como eliminar, actualizar y fusionar a nivel de fila, instrucciones y procedimientos del lenguaje de definición de datos de Spark específicos de Iceberg, como la compactación, la caducidad de las instantáneas, la ramificación y el etiquetado, etc. Consulta lo siguiente para obtener más información:

- Extensiones de escritura de Iceberg Spark: Spark Writes
- Extensiones DDL de Iceberg Spark: extensiones SQL de ALTER TABLE
- Extensiones de procedimientos de Iceberg Spark: Spark Procedures

Consideraciones para el uso de Iceberg con Spark

- De forma predeterminada, Amazon EMR 6.5.0 no admite la ejecución de Iceberg en Amazon EMR en EKS. Hay disponible una imagen personalizada de Amazon EMR 6.5.0 para que pueda pasar `--jars local:///usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` como parámetro `spark-submit` para crear tablas de Iceberg en Amazon EMR en EKS. Para obtener

más información, consulte Enviar una carga de trabajo de Spark en Amazon EMR con una imagen personalizada en la Guía de desarrollo de Amazon EMR en EKS. También puede contactar con Soporte si necesita ayuda. A partir de Amazon EMR 6.6.0, Iceberg es compatible con Amazon EMR en EKS.

- Cuando utilice AWS Glue como catálogo de Iceberg, asegúrese de que la base de datos en la que va a crear la tabla esté en AWS Glue. Si utilizas servicios como estos AWS Lake Formation y no puedes cargar el catálogo, asegúrate de tener el acceso adecuado al servicio para ejecutar el comando.
- Si usa Iceberg SparkSessionCatalog, como se describe enDiferencias de configuración cuando se usa Iceberg SparkCatalog y SparkSessionCatalog, debe seguir los pasos de configuración descritos en Configurar el catálogo de datos de AWS Glue como el metaalmacén de Apache Hive, además de configurar los ajustes del catálogo de datos de Spark Iceberg AWS Glue.

Usar un clúster de Iceberg con Trino

A partir de la versión 6.6.0 de Amazon EMR, puede usar Iceberg con su clúster de Trino.

En este tutorial, utilizará el AWS CLI para trabajar con Iceberg en un clúster Trino de Amazon EMR. Para usar la consola para crear un clúster con Iceberg instalado, siga los pasos que se indican en Crear un lago de datos de Apache Iceberg con Amazon Athena, Amazon EMR y AWS Glue.

Crear un clúster de Iceberg

Para usar Iceberg en Amazon EMR con AWS CLI el, cree primero un clúster con los siguientes pasos. Para obtener información sobre cómo especificar la clasificación de Iceberg mediante la AWS CLI, consulte o. Proporcione una configuración mediante AWS CLI al crear un clúster Proporcione una configuración mediante el SDK de Java al crear un clúster

1. Cree un archivo `iceberg.properties` y establezca un valor para el catálogo que elija. Por ejemplo, si quiere usar el metaalmacén de Hive como catálogo, su archivo debe tener el siguiente contenido.

   ```
   connector.name=iceberg
   hive.metastore.uri=thrift://localhost:9083
   ```

   Si quieres usar el catálogo de datos de AWS Glue como almacén, el archivo debe tener el siguiente contenido.

```
connector.name=iceberg
iceberg.catalog.type=glue
```

2. Cree una acción de arranque que copie `iceberg.properties` desde Amazon S3 a `/etc/trino/conf/catalog/iceberg.properties`, como en el siguiente ejemplo. Para obtener más información sobre las acciones de arranque, consulte Crear acciones de arranque para instalar software adicional.

```
set -ex
sudo aws s3 cp s3://amzn-s3-demo-bucket/iceberg.properties /etc/trino/conf/catalog/
iceberg.properties
```

3. Cree un clúster con la siguiente configuración y sustituya el nombre de la clave y la ruta del script de las acciones de arranque del ejemplo por los suyos propios.

```
aws emr create-cluster --release-label emr-6.7.0 \
--applications Name=Trino \
--region us-east-1 \
--name My_Trino_Iceberg_Cluster \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket","Name":"Add
 iceberg.properties"}]' \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=c3.4xlarge
 InstanceGroupType=CORE,InstanceCount=3,InstanceType=c3.4xlarge \
--use-default-roles \
--ec2-attributes KeyName=<key-name>
```

Inicializar una sesión de Trino para Iceberg

Para iniciar una sesión de Trino, ejecute el siguiente comando.

```
trino-cli --catalog iceberg
```

Escribir en una tabla de Iceberg

Cree la siguiente tabla y escriba en ella con los siguientes comandos de SQL.

```
trino> SHOW SCHEMAS;
trino> CREATE TABLE default.iceberg_table (
            id int,
```

```
            data varchar,
            category varchar)
       WITH (
            format = 'PARQUET',
            partitioning = ARRAY['category', 'bucket(id, 16)'],
            location = 's3://amzn-s3-demo-bucket/<prefix>')

trino> INSERT INTO default.iceberg_table VALUES (1,'a','c1'), (2,'b','c2'),
 (3,'c','c3');
```

Leer desde una tabla para Iceberg

Para leer desde su tabla de Iceberg, ejecute el siguiente comando.

```
trino> SELECT * from default.iceberg_table;
```

Consideraciones para el uso de Iceberg con Trino

- Amazon EMR 6.5 no ofrece compatibilidad nativa con Trino Iceberg Catalog para Iceberg. Trino necesita Iceberg v0.11, por lo que recomendamos lanzar un clúster de Amazon EMR para Trino independiente del clúster de Spark e incluir Iceberg v0.11 en ese clúster.
- Cuando utilice AWS Glue como catálogo para Iceberg, asegúrese de que la base de datos en la que va a crear la tabla existe en AWS Glue. Si utilizas servicios como estos AWS Lake Formation y no puedes cargar el catálogo, asegúrate de tener el acceso adecuado al servicio para ejecutar el comando.
- La integración de Iceberg Glue no funciona con el catálogo de almacenamiento gestionado de Redshift.

Usar un clúster de Iceberg con Flink

A partir de la versión 6.9.0 de Amazon EMR, puede utilizar Iceberg con un clúster de Flink sin los pasos de configuración necesarios al utilizar la integración de código abierto de Flink de Iceberg.

Creación de un clúster de Iceberg

Puede crear un clúster con Iceberg instalado mediante la AWS Management Console, la AWS CLI o la API de Amazon EMR. En este tutorial, utilizará el AWS CLI para trabajar con Iceberg en un clúster de Amazon EMR. Para usar la consola para crear un clúster con Iceberg instalado, siga los pasos

que se indican en Crear un lago de datos de Apache Iceberg con Amazon Athena, Amazon EMR y AWS Glue.

Para usar Iceberg en Amazon EMR con AWS CLI el, cree primero un clúster con los siguientes pasos. Para obtener información sobre cómo especificar la clasificación de Iceberg mediante la AWS CLI, consulte o. Proporcione una configuración mediante AWS CLI al crear un clúster Proporcione una configuración mediante el SDK de Java al crear un clúster Cree un archivo denominado `configurations.json` con el siguiente contenido:

```
[{
"Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

A continuación, cree un clúster con la siguiente configuración y sustituya la ruta del bucket y el ID de subred de Amazon S3 de ejemplo por sus propios valores:

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Flink \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_flink_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Como alternativa, puede crear un clúster de Amazon EMR 6.9.0 con una aplicación de Flink y utilizar el archivo `/usr/share/aws/iceberg/lib/iceberg-flink-runtime.jar` como una dependencia JAR en un trabajo de Flink.

Uso del cliente SQL de Flink

El script del cliente SQL se encuentra en `/usr/lib/flink/bin`. Puede ejecutar el script con el siguiente comando:

```
flink-yarn-session -d # starting the Flink YARN Session in detached mode
./sql-client.sh
```

Esto lanza un intérprete de comandos SQL de Flink.

Ejemplos de Flink

Crear una tabla de Iceberg

SQL de Flink

```
CREATE CATALOG glue_catalog WITH (
   'type'='iceberg',
   'warehouse'='<WAREHOUSE>',
   'catalog-type'='glue'
 );

USE CATALOG  glue_catalog;

CREATE DATABASE IF NOT EXISTS <DB>;

USE <DB>;

CREATE TABLE IF NOT EXISTS `glue_catalog`.`<DB>`.`sample` (id int, data string);
```

API de tabla

```
EnvironmentSettings settings =
                EnvironmentSettings.newInstance().inBatchMode().build();

TableEnvironment tEnv = TableEnvironment.create(settings);

String warehouse = "<WAREHOUSE>";
String db = "<DB>";

tEnv.executeSql(
                "CREATE CATALOG glue_catalog WITH (\n"
                        + "   'type'='iceberg',\n"
                        + "   'warehouse'='"
                        + warehouse
                        + "',\n"
                        + "   'catalog-impl'='org.apache.iceberg.aws.glue.GlueCatalog',
\n"
                        + "   'io-impl'='org.apache.iceberg.aws.s3.S3FileIO'\n"
                        + " );");

tEnv.executeSql("USE CATALOG  glue_catalog;");
tEnv.executeSql("CREATE DATABASE IF NOT EXISTS " + db + ";");
```

```
tEnv.executeSql("USE " + db + ";");
tEnv.executeSql(
        "CREATE TABLE `glue_catalog`.`" + db + "`.`sample` (id bigint, data string);");
```

Escribir en una tabla de Iceberg

SQL de Flink

```
INSERT INTO `glue_catalog`.`<DB>`.`sample` values (1, 'a'),(2,'b'),(3,'c');
```

API de tabla

```
tEnv.executeSql(
        "INSERT INTO `glue_catalog`.`"
                + db
                + "`.`sample` values (1, 'a'),(2,'b'),(3,'c');");
```

API de Datastream

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

GenericRowData rowData1 = new GenericRowData(2);
rowData1.setField(0, 1L);
rowData1.setField(1, StringData.fromString("a"));

DataStream<RowData> input = env.fromElements(rowData1);

Map<String, String> props = new HashMap<();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
```

```
                props,
                new Configuration(),
                "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));

DataStreamSink<Void> dataStreamSink =
        FlinkSink.forRowData(input).tableLoader(tableLoader).append();

env.execute("Datastream Write");
```

Leer desde una tabla de Iceberg

SQL de Flink

```
SELECT * FROM `glue_catalog`.`<DB>`.`sample`;
```

API de tabla

```
Table result = tEnv.sqlQuery("select * from `glue_catalog`.`" + db + "`.`sample`;");
```

API de Datastream

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

Map<String, String> props = new HashMap<>();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
                props,
```

```
                new Configuration(),
                "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));

DataStream<RowData> batch =

 FlinkSource.forRowData().env(env).tableLoader(tableLoader).streaming(false).build();

batch.print().name("print-sink");
```

Uso del catálogo de Hive

Asegúrese de que las dependencias entre Flink y Hive se resuelvan como se describe en Configuración de Flink con el metaalmacén de Hive y el Catálogo de Glue.

Ejecución de un trabajo de Flink

Una forma de enviar un trabajo a Flink es utilizar una sesión de Flink YARN por trabajo. Puede iniciarlo con el siguiente comando:

```
sudo flink run -m yarn-cluster -p 4 -yjm 1024m -ytm 4096m $JAR_FILE_NAME
```

Consideraciones para el uso de Iceberg con Flink

- Cuando utilice AWS Glue como catálogo de Iceberg, asegúrese de que la base de datos en la que va a crear la tabla esté en AWS Glue. Si utilizas servicios como estos AWS Lake Formation y no puedes cargar el catálogo, asegúrate de tener el acceso adecuado al servicio para ejecutar el comando.
- La integración de Iceberg Glue no funciona con el catálogo de almacenamiento gestionado de Redshift.

Usar un clúster de Iceberg con Hive

A partir de la versión 6.9.0 de Amazon EMR, puede utilizar Iceberg con un clúster de Hive sin realizar los pasos de configuración necesarios al utilizar la integración de código abierto de Iceberg Hive. Para las versiones 6.8.0 y anteriores de Amazon EMR, puede utilizar una acción de arranque para instalar el jar `iceberg-hive-runtime` y configurar la compatibilidad de Hive con Iceberg.

Amazon EMR 6.9.0 incluye todas las características de la integración de Hive 3.1.3 con Iceberg 0.14.1 y también incluye características adicionales de Amazon EMR, como la selección automática de los motores de ejecución compatibles en tiempo de ejecución (Amazon EMR en EKS 6.9.0).

Crear un clúster de Iceberg

Puede crear un clúster con Iceberg instalado mediante la AWS Management Console, la AWS CLI o la API de Amazon EMR. En este tutorial, utilizará el AWS CLI para trabajar con Iceberg en un clúster de Amazon EMR. Para usar la consola para crear un clúster con Iceberg instalado, siga los pasos que se indican en Crear un lago de datos de Iceberg con Amazon Athena, Amazon EMR y AWS Glue.

Para usar Iceberg en Amazon EMR con AWS CLI el, cree primero un clúster siguiendo los pasos que se indican a continuación. Para obtener información sobre cómo especificar la clasificación de Iceberg mediante el SDK de Java AWS CLI o el SDK de Java, consulte o. Proporcione una configuración mediante AWS CLI al crear un clúster Proporcione una configuración mediante el SDK de Java al crear un clúster Cree un archivo denominado `configurations.json` con el siguiente contenido:

```
[{
    "Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

A continuación, cree un clúster con la siguiente configuración y sustituya la ruta del bucket y el ID de subred de Amazon S3 de ejemplo por sus propios valores:

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Hive \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_hive_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Un clúster de Hive Iceberg hace lo siguiente:

- Carga en Hive el archivo jar de tiempo de ejecución de Iceberg Hive y habilita la configuración relacionada con Iceberg para el motor Hive.
- Permite la selección del motor de ejecución dinámico de Amazon EMR Hive para evitar que los usuarios configuren el motor de ejecución admitido para que sea compatible con Iceberg.

> **Note**
>
> Los clústeres de Hive Iceberg no son compatibles actualmente con AWS Glue Data Catalog. El catálogo de Iceberg predeterminado es `HiveCatalog`, que corresponde al metaalmacén configurado para el entorno de Hive. Para obtener más información sobre la administración de catálogos, consulte Uso HCatalog en la documentación de Apache Hive.

Compatibilidad de características

Amazon EMR 6.9.0 es compatible con Hive 3.1.3 y con Iceberg 0.14.1. El soporte de características se limita a las características compatibles con Iceberg para Hive 3.1.2 y 3.1.3. Se admiten los siguientes comandos:

- Con las versiones 6.9.0 a 6.12.x de Amazon EMR, debe incluir el archivo `libfb303` jar en el directorio `auxlib` de Hive. Ejecute el siguiente comando para incluirlo:

  ```
  sudo /usr/bin/ln -sf /usr/lib/hive/lib/libfb303-*.jar /usr/lib/hive/auxlib/
  libfb303.jar
  ```

  Con las versiones 6.13 y posteriores de Amazon EMR, el archivo `libfb303` jar se vincula de forma simbólica automáticamente al directorio `auxlib` de Hive.
- Creación de una tabla
  - Tabla no particionada: se pueden crear tablas externas en Hive proporcionando el siguiente controlador de almacenamiento:

    ```
    CREATE EXTERNAL TABLE x (i int) STORED BY
     'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
    ```
  - Tabla particionada: las tablas particionadas externas en Hive se pueden crear de la siguiente manera:

```
CREATE EXTERNAL TABLE x (i int) PARTITIONED BY (j int) STORED BY
 'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
```

> **Note**
>
> El formato de STORED AS archivo de no ORC/AVRO/PARQUET es compatible con Hive 3. La única opción por defecto es Parquet.

- Eliminación de una tabla: el comando DROP TABLE se usa para eliminar tablas, como en el siguiente ejemplo:

```
DROP TABLE [IF EXISTS] table_name [PURGE];
```

- Lectura de una tabla: las instrucciones SELECT se pueden usar para leer tablas de Iceberg en Hive, como en el siguiente ejemplo. Los motores de ejecución compatibles son MR y Tez.

```
SELECT * FROM table_name
```

  Para obtener información sobre la sintaxis de selección de Hive, consulte LanguageManual Select. Para obtener información sobre determinadas instrucciones con tablas de Iceberg en Hive, consulte Selección de Apache Iceberg.

- Inserción en una tabla: la instrucción INSERT INTO de HiveQL solo funciona en tablas de Iceberg y solo es compatible con el motor de ejecución Map Reduce. Los usuarios de Amazon EMR no necesitan configurar de forma explícita el motor de ejecución porque Amazon EMR Hive selecciona el motor para tablas de Iceberg en tiempo de ejecución.
  - Inserción de una sola tabla: ejemplo:

```
INSERT INTO table_name VALUES ('a', 1);
INSERT INTO table_name SELECT...;
```

  - Inserción de varias tablas: se admite la inserción de varias tablas no atómicas en instrucciones. Ejemplo:

```
FROM source
 INSERT INTO table_1 SELECT a, b
 INSERT INTO table_2 SELECT c,d;
```

A partir de Amazon EMR 7.3.0, Hive with Iceberg admite el catálogo de datos de AWS Glue como metaalmacén. Para usar el catálogo de datos de AWS Glue como metaalmacén, defina la siguiente propiedad.

```
SET iceberg.catalog.<catalog_name>.catalog-
impl=org.apache.iceberg.aws.glue.GlueCatalog;
```

Si lo desea, también puede especificar la siguiente propiedad.

```
SET iceberg.catalog.<catalog_name>.type=glue;
```

Puede crear una tabla utilizando el siguiente ejemplo.

```
CREATE EXTERNAL TABLE table_name (col1 type1, col2 type2,..)
ROW FORMAT SERDE 'org.apache.iceberg.mr.hive.HiveIcebergSerDe'
STORED BY 'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
location '<location>'
TBLPROPERTIES ('table_type'='iceberg', 'iceberg.catalog'='<catalog_name>');
```

Consideraciones para el uso de Iceberg con Hive

- Iceberg admite los siguientes tipos de consultas:
  - Crear tablas
  - Eliminar tablas
  - Insertar en tabla
  - Leer tabla
- Solo el motor de ejecución MR (MapReduce) es compatible con las operaciones de DML (lenguaje de manipulación de datos), y el MR está obsoleto en Hive 3.1.3.
- AWS El catálogo de datos de Glue no es compatible actualmente con Iceberg with Hive.
- La gestión de errores no es lo suficientemente sólida. En caso de configuración incorrecta, es posible que las inserciones en las consultas se completen correctamente. Sin embargo, la no actualización de los metadatos puede provocar la pérdida de datos.

- La integración de Iceberg Glue no funciona con el catálogo de almacenamiento gestionado de Redshift.

Historial de versiones de Iceberg

En la siguiente tabla, se muestra la versión de Iceberg incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Iceberg

Etiqueta de versión de Amazon EMR	Versión de Iceberg	Componentes instalados con Iceberg
emr-7.9.0	1.7.1-amzn-2	Not available.
emr-7.8.0	1.7.1-amzn-1	Not available.
emr-7.7.0	1.7.1-amzn-0	Not available.
emr-7.6.0	1.6.1-amzn-2	Not available.
emr-7.5.0	1.6.1-amzn-1	Not available.
emr-7.4.0	1.6.1-amzn-0	Not available.
emr-7.3.0	1.5.2-amzn-0	Not available.
emr-7.2.0	1.5.0-amzn-0	Not available.
emr-7.1.0	1.4.3-amzn-0	Not available.
emr-7.0.0	1.4.2-amzn-0	Not available.
emr-6.15.0	1.4.0-amzn-0	Not available.
emr-6.14.0	1.3.1-amzn-0	Not available.
emr-6.13.0	1.3.0-amzn-1	Not available.

Etiqueta de versión de Amazon EMR	Versión de Iceberg	Componentes instalados con Iceberg
emr-6.12.0	1.3.0-amzn-0	Not available.
emr-6.11.1	1.2.0-amzn-0	Not available.
emr-6.11.0	1.2.0-amzn-0	Not available.
emr-6.10.1	1.1.0-amzn-0	Not available.
emr-6.10.0	1.1.0-amzn-0	Not available.
emr-6.9.1	0.14.1-amzn-0	Not available.
emr-6.9.0	0.14.1-amzn-0	Not available.
emr-6.8.1	0.14.0-amzn-0	Not available.
emr-6.8.0	0.14.0-amzn-0	Not available.
emr-6.7.0	0.13.1-amzn-0	Not available.
emr-6.6.0	0.13.1	Not available.
emr-6.5.0	0.12.0	Not available.

Notas de la versión de Iceberg por versión

- Amazon EMR 6.9.0: notas de la versión de Iceberg

Amazon EMR 6.9.0: notas de la versión de Iceberg

Amazon EMR 6.9.0: cambios en Iceberg

Tipo	Descripción
Característica	Integración de Amazon EMR Flink con Iceberg.

Tipo	Descripción
Característica	Integración de Amazon EMR Hive con Iceberg.
Característica	Support para almacenar en caché los archivos de metadatos de Iceberg en Amazon FSx for Lustre para mejorar el tiempo de planificación de las consultas.
Portabilidad con versiones anteriores	PR 5050: Flink 1.15: compatibilidad de opciones de escritura en comentarios SQL de inserción en línea.
Portabilidad con versiones anteriores	PR 5282: AWS: Corrija los errores de reintento de PUT abriendo nuevos flujos de archivos de datos.
Portabilidad con versiones anteriores	PR 5318: Flink 1.15: cierra la brecha entre FlinkSource y IcebergSource (FLIP-27) y se agregó una configuración opcional para usar el código fuente FLIP-27 en Flink SQL.
Portabilidad con versiones anteriores	PR 5344: Flink 1.14: cierra la brecha entre FlinkSource y IcebergSource (FLIP-27) y añade una configuración opcional para usar el código fuente FLIP-27 en Flink SQL.
Portabilidad con versiones anteriores	PR 5393: Flink 1.14, 1.15: Evite convertir las métricas de Iceberg en Flink en el lector de código fuente FLIP-27. MetricContext
Portabilidad con versiones anteriores	PR 5401: Flink 1.14, 1.15: Se omitió un grupo en el PR #5393 para las métricas del lector de código fuente FLIP-27. IcebergSourceReader
Portabilidad con versiones anteriores	PR 5679: Spark 3.2, 3.3: Se corrige la propagación de la nulabilidad en el nodo. MergeRows

Tipo	Descripción
Portabilidad con versiones anteriores	PR 5860: Spark 3.3: Se corrige QueryFailure cuando se ejecuta RewriteManifestProcedure en tablas particionadas por fechas.
Portabilidad con versiones anteriores	PR 5880: Spark 3.3: Se corrige la nulabilidad en las proyecciones. merge-on-read
Portabilidad con versiones anteriores	PR 5917: Spark 3.2: Se corrige la nulabilidad en las proyecciones. merge-on-read

Cuaderno de Jupyter en Amazon EMR

Cuaderno de Jupyter es una aplicación web de código abierto que puede utilizar para crear y compartir documentos que contienen código activo, ecuaciones, visualizaciones y texto narrativo. Amazon EMR ofrece dos opciones para trabajar con los cuadernos de Jupyter:

Temas

- EMR Studio
- Cuadernos de Amazon EMR basados en cuaderno de Jupyter
- JupyterHub

EMR Studio

Amazon EMR Studio es un entorno de desarrollo integrado (IDE) basado en web para cuadernos de Jupyter completamente administrados que se ejecutan en clústeres de Amazon EMR. Puede configurar un estudio de EMR para que su equipo desarrolle, visualice y depure aplicaciones escritas en R, Python, Scala y. PySpark

Recomendamos utilizar EMR Studio cuando utilice cuadernos de Jupyter en Amazon EMR. Para más información, consulte EMR Studio en la Guía de administración de Amazon EMR.

Cuadernos de Amazon EMR basados en cuaderno de Jupyter

Cuadernos de EMR es un entorno de cuaderno de Jupyter integrado en la consola Amazon EMR que le permite crear rápidamente cuadernos de Jupyter, adjuntarlos a los clústeres de Spark y, a continuación, abrir el editor del cuaderno de Jupyter en la consola para ejecutar consultas y código de forma remota. Un cuaderno de EMR se guarda en Amazon S3 de forma independiente de los clústeres para un almacenamiento duradero, un acceso rápido y flexibilidad. Puede tener varios cuadernos abiertas, adjuntar varios cuadernos a un único clúster y reutilizar un cuaderno en distintos clústeres.

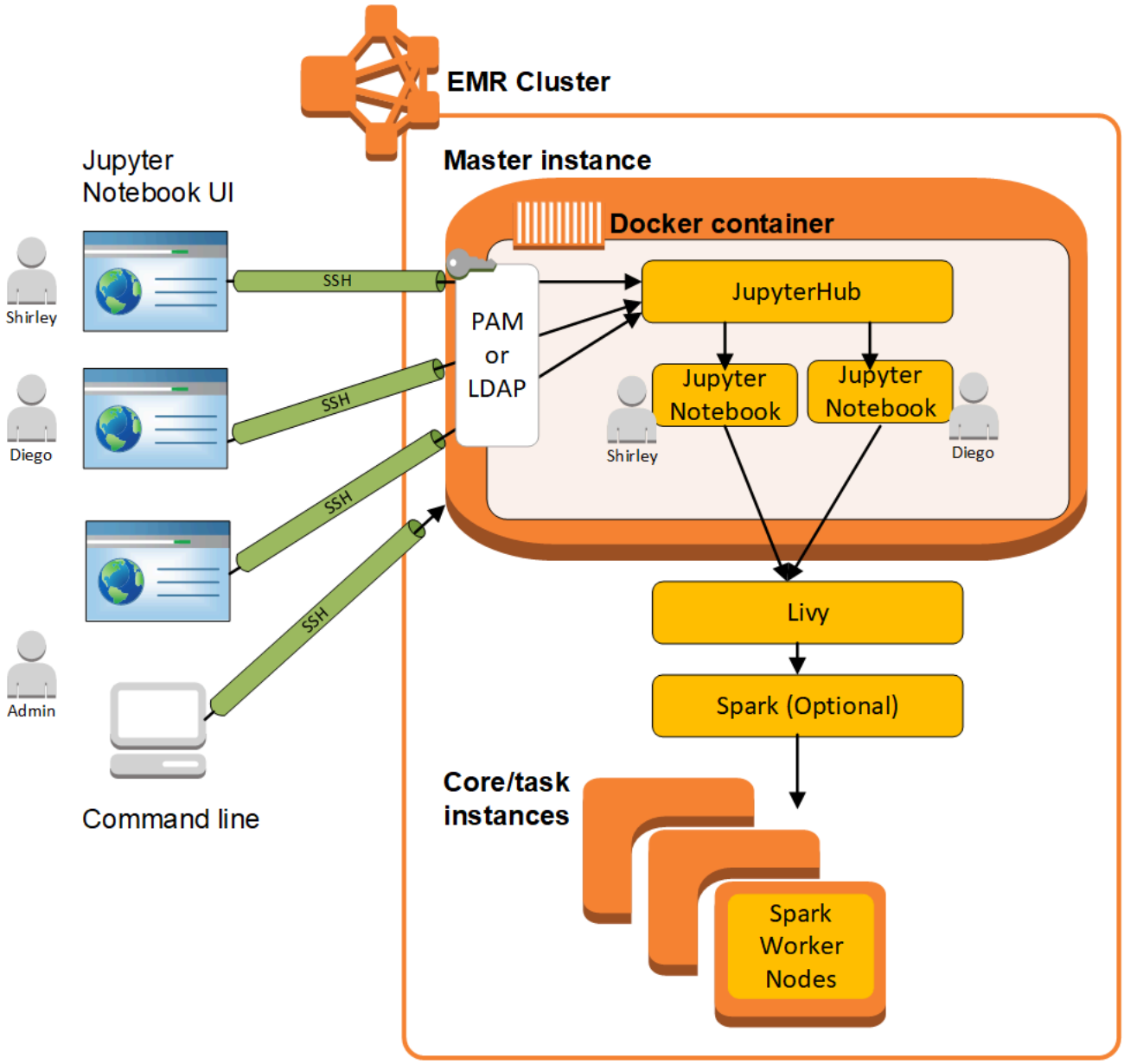
Para más información, consulte Cuadernos de EMR en la Guía de administración de Amazon EMR.

JupyterHub

Cuaderno de Jupyter es una aplicación web de código abierto que puede utilizar para crear y compartir documentos que contienen código activo, ecuaciones, visualizaciones y texto narrativo. JupyterHuble permite alojar múltiples instancias de un servidor de cuadernos de Jupyter de un único usuario. Cuando se crea un clúster con JupyterHub, Amazon EMR crea un contenedor de Docker en el nodo maestro del clúster. JupyterHub, todos los componentes necesarios para Jupyter y Sparkmagic se ejecutan en el contenedor.

Sparkmagic es una biblioteca de kernels que permite que los blocs de notas de Jupyter interactúen con Apache Spark que se ejecuta en Amazon EMR a través de Apache Livy, que es un servidor de REST para Spark. Spark y Apache Livy se instalan automáticamente al crear un clúster con JupyterHub. El kernel de Python 3 predeterminado para Jupyter está disponible junto con los kernels PySpark 3 PySpark, y Spark que están disponibles con Sparkmagic. Puede utilizar estos kernels para ejecutar código Spark ad hoc y consultas SQL interactiva con Python y Scala. Puede instalar kernels adicionales en el contenedor de Docker manualmente. Para obtener más información, consulte Instalación de kernels y bibliotecas adicionales.

En el siguiente diagrama, se muestran los componentes de JupyterHub Amazon EMR con los métodos de autenticación correspondientes para los usuarios de los blocs de notas y para el administrador. Para obtener más información, consulte Agregar usuarios y administradores de Cuadernos de Jupyter.



En la tabla siguiente, se muestra la versión de JupyterHub incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes con los que Amazon EMR se instala. JupyterHub

Para ver la versión de los componentes instalados con JupyterHub esta versión, consulte Versiones de los componentes de la versión 7.9.0.

JupyterHub información de la versión de emr-7.9.0

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr7.9.0	JupyterHub 1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

En la tabla siguiente, se muestra la versión de JupyterHub incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes con los que Amazon EMR se instala. JupyterHub

Para ver la versión de los componentes instalados con JupyterHub esta versión, consulte Versiones del componente de la versión 6.15.0.

JupyterHub información de la versión de emr-6.15.0

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.15.0	JupyterHub 1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
		, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

En la tabla siguiente, se muestra la versión de JupyterHub incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes con los componentes con los que Amazon EMR se instala. JupyterHub

Para ver la versión de los componentes instalados con JupyterHub esta versión, consulte Versiones del componente de la versión 5.36.2.

JupyterHub información de la versión de emr-5.36.2

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.36.2	JupyterHub 1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

El kernel de Python 3 que se incluye JupyterHub en Amazon EMR es la versión 3.6.4.

Las bibliotecas instaladas en el `jupyterhub` contenedor pueden variar según las versiones de lanzamiento de Amazon EMR y las versiones de AMI de Amazon EC2 AMI.

Para enumerar las bibliotecas instaladas utilizando **`conda`**

- Ejecute el siguiente comando en la línea de comandos del nodo principal:

```
sudo docker exec jupyterhub bash -c "conda list"
```

Para enumerar las bibliotecas instaladas utilizando **`pip`**

- Ejecute el siguiente comando en la línea de comandos del nodo principal:

```
sudo docker exec jupyterhub bash -c "pip freeze"
```

Temas

- Cree un clúster con JupyterHub
- Consideraciones sobre el uso JupyterHub en Amazon EMR
- Configurando JupyterHub
- Configuración de la persistencia de los cuadernos en Amazon S3
- Conexión al nodo maestro y a los servidores de blocs de notas
- JupyterHub configuración y administración
- Agregar usuarios y administradores de Cuadernos de Jupyter
- Instalación de kernels y bibliotecas adicionales
- JupyterHub historial de lanzamientos

Cree un clúster con JupyterHub

Puede crear un clúster de Amazon EMR JupyterHub mediante la AWS Management Console AWS Command Line Interface, la o la API de Amazon EMR. Asegúrese de que el clúster no se crea con la opción de terminar automáticamente después de completar los pasos (opción `--auto-terminate` de la AWS CLI). Además, asegúrese de que los administradores y los usuarios de blocs de notas puedan obtener acceso al par de claves que utiliza al crear el clúster. Para obtener más información,

consulte Usar un par de claves para credenciales SSH en la Guía de administración de Amazon EMR.

Crear un clúster JupyterHub con la consola

Utilice el siguiente procedimiento para crear un clúster JupyterHub instalado mediante Opciones avanzadas en la consola de Amazon EMR.

Para crear un clúster de Amazon EMR JupyterHub instalado mediante la consola de Amazon EMR

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. En Software Configuration (Configuración de software):
   - En Versión, seleccione emr-5.36.2 y elija. JupyterHub
   - Si utiliza Spark, para usar el Catálogo de datos de AWS Glue como metaalmacén de Spark SQL, seleccione Usar para metadatos de la tabla de Spark. Para obtener más información, consulte Usa el catálogo de AWS Glue Data Catalog con Spark en Amazon EMR.
   - En Edit software settings (Editar configuración de software) elija Enter configuration (Escribir la configuración) y especifique los valores o elija Load JSON from S3 (Cargar JSON desde S3) y especifique un archivo de configuración JSON. Para obtener más información, consulte Configurando JupyterHub.
4. En Add steps (optional) (Añadir pasos (opcional)) configure los pasos que se ejecutarán cuando se cree el clúster, asegúrese de que no se ha seleccionado Auto-terminate cluster after the last step is completed (Terminar automáticamente el clúster después de que se complete el último paso) y elija Next (Siguiente).
5. Elija opciones de Hardware Configuration (Configuración de hardware) y Next (Siguiente). Para obtener más información, consulte Configuración del hardware y las redes de los clústeres en la Guía de administración de Amazon EMR.
6. Elija las opciones de General Cluster Settings (Configuración general del clúster) y Next (Siguiente).
7. Elija Security Options (Opciones de seguridad), especifique un par de claves y, a continuación, elija Create Cluster (Crear clúster).

Crea un clúster JupyterHub con el AWS CLI

Para lanzar un clúster con JupyterHub, utilice el `aws emr create-cluster` comando y, para la `--applications` opción, especifique`Name=JupyterHub`. En el siguiente ejemplo, se lanza un JupyterHub clúster en Amazon EMR con dos EC2 instancias (una principal y otra secundaria). Además, la depuración está habilitada y los registros se almacenan en la ubicación de Amazon S3 especificada mediante `--log-uri`. El par de claves especificado proporciona acceso a EC2 las instancias de Amazon del clúster.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair
```

Consideraciones sobre el uso JupyterHub en Amazon EMR

Cuando utilice Amazon EMR, tenga JupyterHub en cuenta lo siguiente.

- > **Warning**
  >
  > Los blocs de notas y los archivos de usuario se guardan en el sistema de archivos en el nodo principal. Se trata de un almacenamiento efímero que no se conserva al terminar el clúster. Cuando un clúster termina, estos datos se pierden si no se ha hecho una copia de seguridad. Le recomendamos que programe copias de seguridad periódicas con trabajos `cron` u otros medios adecuados para su aplicación.
  > Además, los cambios de configuración realizados en el contenedor no se pueden mantener si se reinicia el contenedor. Le recomendamos que cree un script o automatice la configuración de contenedores de otra forma para que pueda reproducir personalizaciones más fácilmente.

- No se admite la autenticación de Kerberos que se ha configurado con una configuración de seguridad de Amazon EMR.
- OAuthenticator no se admite.

Configurando JupyterHub

Puede personalizar la configuración de Amazon EMR y de JupyterHub los cuadernos de usuarios individuales si se conecta al nodo maestro del clúster y edita los archivos de configuración. Después de cambiar los valores, reinicie el contenedor `jupyterhub`.

Modifique las propiedades de los siguientes archivos para configurar los cuadernos JupyterHub de Jupyter individuales:

- `jupyterhub_config.py`: de forma predeterminada, este archivo se guarda en el directorio de `/etc/jupyter/conf/` del nodo maestro. Para obtener más información, consulte Conceptos básicos de la configuración en la JupyterHub documentación.
- `jupyter_notebook_config.py`: este archivo se guarda en el directorio `/etc/jupyter/` de forma predeterminada y se copia en el contenedor `jupyterhub` como la opción predeterminada. Para obtener más información, consulte Config file and command line options en la documentación de Jupyter Notebook.

También puede utilizar la clasificación de configuración `jupyter-sparkmagic-conf` al crear un clúster para personalizar Sparkmagic, que actualiza los valores del archivo `config.json` para Sparkmagic. Para obtener más información sobre los ajustes disponibles, consulta example_config.json en. GitHub Para obtener más información sobre cómo usar clasificaciones de configuración con aplicaciones en Amazon EMR, consulte Configuración de aplicaciones.

En el siguiente ejemplo, se lanza un clúster utilizando el archivo `MyJupyterConfig.json` para los ajustes AWS CLI de clasificación de la configuración de Sparkmagic.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --use-default-roles --release-label emr-5.14.0 \
--applications Name=Jupyter --instance-type m4.xlarge --instance-count 3 \
--ec2-attributes KeyName=MyKey,SubnetId=subnet-1234a5b6 --configurations file://
MyJupyterConfig.json
```

El contenido de ejemplo de `MyJupyterConfig.json` es el siguiente:

```
[
    {
    "Classification":"jupyter-sparkmagic-conf",
    "Properties": {
      "kernel_python_credentials" : "{\"username\":\"diego\",\"base64_password\":
\"mypass\",\"url\":\"http:\/\/localhost:8998\",\"auth\":\"None\"}"
      }
    }
]
```

> **Note**
>
> Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Configuración de la persistencia de los cuadernos en Amazon S3

Puede configurar un JupyterHub clúster en Amazon EMR de forma que los cuadernos guardados por un usuario se conserven Amazon S3, fuera del almacenamiento efímero de las instancias de clúster. EC2

Para especificar la persistencia en Amazon S3, utilice la clasificación de configuración `jupyter-s3-conf` al crear el clúster. Para obtener más información, consulte Configuración de aplicaciones.

Además de habilitar la persistencia de Amazon S3 utilizando la propiedad `s3.persistence.enabled`, debe utilizar la propiedad `s3.persistence.bucket` para especificar el bucket en Amazon S3 donde se guardan los cuadernos. Los blocs de notas de cada usuario se guardan en una carpeta `jupyter/`*`jupyterhub-user-name`* del bucket

especificado. El bucket ya debe existir en Amazon S3, y el rol para el perfil de EC2 instancia que especifique al crear el clúster debe tener permisos para el bucket (de forma predeterminada, el rol es`EMR_EC2_DefaultRole`). Para obtener más información, consulte Configuración de roles de IAM para permisos AWS de Amazon EMR para los servicios de.

Cuando se lanza un nuevo clúster con las mismas propiedades de clasificación de configuración, los usuarios pueden abrir los blocs de notas con el contenido de la ubicación donde se guardaron.

Tenga en cuenta que, si importa archivos como módulos en un cuaderno con Amazon S3 activado, los archivos se subirán a Amazon S3. Cuando importa archivos sin habilitar la persistencia de Amazon S3, se cargan en su JupyterHub contenedor.

En el siguiente ejemplo, se habilita la persistencia de Amazon S3. Los blocs de notas guardados por los usuarios se almacenan en la carpeta `s3://MyJupyterBackups/jupyter/`*`jupyterhub-user-name`* de cada usuario, donde *`jupyterhub-user-name`* es un nombre de usuario, como por ejemplo `diego`.

```
[
    {
        "Classification": "jupyter-s3-conf",
        "Properties": {
            "s3.persistence.enabled": "true",
            "s3.persistence.bucket": "MyJupyterBackups"
        }
    }
]
```

Conexión al nodo maestro y a los servidores de blocs de notas

JupyterHub los administradores y los usuarios de portátiles deben conectarse al nodo principal del clúster mediante un túnel SSH y, a continuación, conectarse a las interfaces web servidas por JupyterHub el nodo maestro. Para obtener más información sobre la configuración de un túnel SSH y el uso del túnel para conexiones web proxy, consulte Conectar con el clúster en la Guía de administración de Amazon EMR.

De forma predeterminada, JupyterHub en Amazon EMR está disponible a través del puerto 9443 en el nodo maestro. El JupyterHub proxy interno también sirve a las instancias de notebook a través del puerto 9443. JupyterHub y se puede acceder a las interfaces web de Jupyter mediante una URL con el siguiente patrón:

https://:9443 *MasterNodeDNS*

Puede especificar otro puerto con la propiedad `c.JupyterHub.port` en el archivo `jupyterhub_config.py`. Para obtener más información, consulte Conceptos básicos de redes en la documentación. JupyterHub

De forma predeterminada, JupyterHub en Amazon EMR utiliza un certificado autofirmado para el cifrado SSL mediante HTTPS. Se pide a los usuarios que confíen en el certificado autofirmado cuando se conecten. Puede utilizar un certificado de confianza y sus propias claves. Reemplace el archivo de certificado predeterminado, `server.crt`, y el archivo de claves `server.key` en el directorio `/etc/jupyter/conf/` del nodo principal por sus propios archivos de certificados y claves. Utilice las propiedades `c.JupyterHub.ssl_key` y `c.JupyterHub.ssl_cert` en el archivo `jupyterhub_config.py` para especificar sus materiales SSL. Para obtener más información, consulte la configuración de seguridad en la JupyterHub documentación. Después de actualizar `jupyterhub_config.py`, reinicie el contenedor.

JupyterHub configuración y administración

JupyterHub y los componentes relacionados se ejecutan dentro de un contenedor Docker cuyo nombre `jupyterhub` es el que ejecuta el sistema operativo Ubuntu. Existen varias maneras en las que puede administrar los componentes que se ejecutan en el contenedor.

> **Warning**
>
> Es posible que las personalizaciones que realice en el contenedor no se mantengan si se reinicia el contenedor. Le recomendamos que cree un script o automatice la configuración de contenedores de otra forma para que pueda reproducir personalizaciones más fácilmente.

Administración mediante la línea de comandos

Cuando se conecta al nodo principal mediante SSH, puede emitir comandos utilizando la interfaz de línea de comandos (CLI) de Docker y especificando el contenedor por nombre (`jupyterhub`) o ID. Por ejemplo, `sudo docker exec jupyterhub` *`command`* ejecuta los comandos que reconoce el sistema operativo o una aplicación que se ejecuta en el contenedor. Puede utilizar este método para añadir usuarios al sistema operativo y para instalar aplicaciones y bibliotecas adicionales en el contenedor de Docker. Por ejemplo, la imagen de contenedor predeterminada incluye Conda para la instalación de paquetes, por lo que puede ejecutar el siguiente comando en la línea de comandos del nodo principal para instalar una aplicación, Keras, en el contenedor:

```
sudo docker exec jupyterhub conda install keras
```

Administración mediante el envío de pasos

Los pasos son una forma de enviar el trabajo a un clúster. Puede enviar pasos cuando al lanzar un clúster o bien puede enviar pasos a un clúster en ejecución. Los comandos que se ejecutan en la línea de comandos se pueden enviar como pasos usando `command-runner.jar`. Para obtener más información, consulte Uso de pasos con la CLI y la consola en la Guía de administración de Amazon EMR y Ejecución de comandos y scripts en un clúster de Amazon EMR.

Por ejemplo, puede usar el siguiente AWS CLI comando en un equipo local para instalar Keras de la misma manera que lo hizo desde la línea de comandos del nodo maestro en el ejemplo anterior:

```
aws emr add-steps --cluster-id MyClusterID --steps Name="Command
 Runner",Jar="command-runner.jar",Args="/usr/bin/sudo","/usr/bin/
docker","exec","jupyterhub","conda","install","keras"
```

También puede crear un script de una secuencia de pasos, cargar el script en Amazon S3 y usar `script-runner.jar` para ejecutar el script cuando cree el clúster o agregar el script como un paso. Para obtener más información, consulte Ejecución de comandos y scripts en un clúster de Amazon EMR. Para ver un ejemplo, consulta the section called “Ejemplo: script bash para agregar varios usuarios”.

Administración mediante REST APIs

Jupyter y el proxy HTTP para JupyterHub proporcionar REST APIs que puede usar para enviar solicitudes. JupyterHub Para enviar solicitudes a JupyterHub, debes pasar un token de API con la solicitud. Puede utilizar el comando `curl` de la línea de comandos del nodo principal para ejecutar comandos REST. Para obtener más información, consulte los siguientes recursos:

- Utiliza JupyterHub la API REST en la documentación de JupyterHub, que incluye instrucciones para generar los tokens de API
- API del servidor de Cuaderno de Jupyter en GitHub
- configurable-http-proxyactivado GitHub

En el siguiente ejemplo, se muestra el uso de la API REST JupyterHub para obtener una lista de usuarios. El comando pasa un token de administración generado previamente y usa el puerto predeterminado, 9443 JupyterHub, para canalizar el resultado a jq para facilitar la visualización:

```
curl -XGET -s -k https://$HOST:9443/hub/api/users \
-H "Authorization: token $admin_token" | jq .
```

Agregar usuarios y administradores de Cuadernos de Jupyter

Puede usar uno de dos métodos para que los usuarios se autentiquen con el fin de JupyterHub poder crear blocs de notas y, opcionalmente, administrar. JupyterHub El método más sencillo es utilizar el módulo JupyterHub de autenticación conectable (PAM). Además, JupyterHub en Amazon EMR admite el complemento de autenticación LDAP para JJD para JJM JupyterHub para JJD admite el complemento de autenticación LDAP para JD para JD. En esta sección se proporcionan instrucciones y ejemplos para añadir usuarios con cada método de autenticación.

JupyterHub en Amazon EMR tiene un usuario predeterminado con permisos de administrador. El nombre de usuario es `jovyan` y la contraseña es `jupyter`. Le recomendamos que reemplace el usuario por otro que tenga permisos administrativos. Para ello, puede utilizar un paso cuando cree el clúster o conectarse al nodo principal cuando se ejecute el clúster.

Temas

- Uso de la autenticación PAM
- Uso de la autenticación LDAP
- Suplantación de usuarios

Uso de la autenticación PAM

El proceso de crear usuarios de PAM JupyterHub en Amazon EMR consta de dos pasos. El primer paso es añadir usuarios al sistema operativo que se está ejecutando en el contenedor `jupyterhub` del nodo principal y añadir un directorio de inicio correspondiente a cada usuario. El segundo paso es agregar estos usuarios del sistema operativo como JupyterHub usuarios, un proceso conocido como creación de lista de permitidos. JupyterHub Una vez agregado un JupyterHub usuario, este puede conectarse a la JupyterHub URL y proporcionar las credenciales de acceso de su sistema operativo.

Cuando un usuario inicia sesión, JupyterHub abre la instancia del servidor portátil de ese usuario, que se guarda en el directorio principal del usuario, en el nodo principal, que es`/var/lib/jupyter/home/`*`username`*. Si no existe ninguna instancia de servidor de cuadernos, JupyterHub genera una instancia de cuaderno en el directorio de inicio del usuario. En las siguientes secciones

se muestra cómo añadir usuarios de forma individual al sistema operativo y a él JupyterHub, seguido de un rudimentario script bash que añade varios usuarios.

Agregar un usuario de sistema operativo al contenedor

El siguiente ejemplo primero usa el comando useradd en el contenedor para añadir un solo usuario, diego, y crear un directorio de inicio para ese usuario. El segundo comando utiliza chpasswd para establecer la contraseña diego para este usuario. Los comandos se ejecutan en la línea de comandos del nodo principal mientras están conectados mediante SSH. También puede ejecutar estos comandos usando un paso tal como se ha descrito anteriormente en Administración mediante el envío de pasos.

```
sudo docker exec jupyterhub useradd -m -s /bin/bash -N diego
sudo docker exec jupyterhub bash -c "echo diego:diego | chpasswd"
```

Añadir un usuario JupyterHub

Puede usar el panel de administración JupyterHub o la API REST para agregar usuarios y administradores, o solo usuarios.

Para agregar usuarios y administradores usando el panel de administración en JupyterHub

1. Conéctese al nodo principal mediante SSH e inicie sesión en https: //:9443 *MasterNodeDNS* con una identidad que tenga permisos de administrador.
2. Elija Control Panel (Panel de control), Admin (Administración).
3. Elija User (Usuario), Add Users (Añadir usuarios) o elija Admin (Administrador), Add Admins (Añadir administradores).

Para añadir un usuario con la API de REST

1. Conéctese al nodo principal con SSH y utilice el siguiente comando en él o ejecute el comando como un paso.
2. Adquiera un token administrativo para realizar solicitudes a la API y sustitúyalo por ese token *AdminToken* en el siguiente paso.
3. Usa el siguiente comando y *UserName* sustitúyelo por un usuario del sistema operativo que se haya creado en el contenedor.

```
curl -XPOST -H "Authorization: token AdminToken" "https://$(hostname):9443/hub/api/
users/UserName
```

> **Note**
>
> Se le agrega automáticamente como usuario JupyterHub no administrador cuando inicia sesión en la interfaz JupyterHub web por primera vez.

Ejemplo: script bash para agregar varios usuarios

El siguiente ejemplo de script bash combina los pasos anteriores de esta sección para crear varios JupyterHub usuarios. El script puede ejecutarse directamente en el nodo maestro o puede cargarse en Amazon S3 y, a continuación, ejecutarse como un paso.

El script establece primero una matriz de nombres de usuario y utiliza el comando `jupyterhub token` para crear un token de API para el administrador predeterminado, jovyan. A continuación, crea un usuario de sistema operativo en el contenedor `jupyterhub` para cada usuario, asignando a cada uno una contraseña inicial igual que su nombre de usuario. Por último, llama a la operación de la API REST para crear a cada usuario. JupyterHub Pasa el token generado anteriormente en el script y canaliza la respuesta REST a `jq` para facilitar su visualización.

```
# Bulk add users to container and JupyterHub with temp password of username
set -x
USERS=(shirley diego ana richard li john mary anaya)
TOKEN=$(sudo docker exec jupyterhub /opt/conda/bin/jupyterhub token jovyan | tail -1)
for i in "${USERS[@]}";
do
   sudo docker exec jupyterhub useradd -m -s /bin/bash -N $i
   sudo docker exec jupyterhub bash -c "echo $i:$i | chpasswd"
   curl -XPOST --silent -k https://$(hostname):9443/hub/api/users/$i \
 -H "Authorization: token $TOKEN" | jq
done
```

Guarde el script en una ubicación en Amazon S3, como `s3://amzn-s3-demo-bucket/createjupyterusers.sh`. A continuación, puede utilizar `script-runner.jar` para ejecutarlo como un paso.

Ejemplo: ejecución del script al crear un clúster (AWS CLI)

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair \
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-
s3-demo-bucket/createjupyterusers.sh"]
```

Ejecución del script en un clúster existente (AWS CLI)

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr add-steps --cluster-id j-XXXXXXXX --steps Type=CUSTOM_JAR,\
Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-
s3-demo-bucket/createjupyterusers.sh"]
```

Uso de la autenticación LDAP

El protocolo ligero de acceso a directorios (LDAP) es un protocolo de aplicación para consultar y modificar objetos que corresponden a recursos como usuarios y equipos almacenados en un proveedor de servicios de directorio compatible con LDAP, como Active Directory o un servidor OpenLDAP. Puede utilizar el complemento de autenticación LDAP para JupyterHub Amazon EMR

para JupyterHub usar LDAP para la autenticación de usuarios. El complemento controla las sesiones de inicio de sesión para usuarios LDAP y proporciona información de usuario a Jupyter. Esto permite a los usuarios conectarse a sus libretas JupyterHub y cuadernos utilizando las credenciales de sus identidades almacenadas en un servidor compatible con LDAP.

En los pasos de esta sección se explica cómo configurar y habilitar LDAP mediante el complemento de autenticación de LDAP para. JupyterHub Los pasos se realizan mientras está conectado a la línea de comandos del nodo principal. Para obtener más información, consulte Conexión al nodo maestro y a los servidores de blocs de notas.

1. Cree un archivo de configuración LDAP con información sobre el servidor LDAP, como la dirección IP del host, el puerto, los nombres de enlace, etc.
2. Modifique `/etc/jupyter/conf/jupyterhub_config.py` para habilitar el complemento de autenticación LDAP para JupyterHub.
3. Cree y ejecute un script que configure LDAP en el contenedor `jupyterhub`.
4. Consulte el LDAP para los usuarios y, a continuación, cree los directorios principales dentro del contenedor para cada usuario. JupyterHub requiere que los directorios principales alojen los blocs de notas.
5. Ejecute un script que se reinicie JupyterHub

> **Important**
>
> Antes de configurar LDAP, pruebe la infraestructura de red para asegurarse de que el servidor LDAP y el nodo principal del clúster puedan comunicarse según sea necesario. TLS suele utilizar el puerto 389 a través de una conexión TCP sencilla. Si su conexión LDAP utiliza SSL, el puerto TCP conocido para SSL es 636.

Crear del archivo de configuración de LDAP

En el siguiente ejemplo se utilizan los siguientes valores de configuración de marcador de posición. Reemplácelos por parámetros que coincidan con su implementación.

- El servidor LDAP ejecuta la versión 3 y está disponible en el puerto 389. Es el puerto no SSL estándar para LDAP.
- El nombre distinguido (DN) base es `dc=example, dc=org`.

Utilice un editor de texto para crear el archivo ldap.conf, con un contenido similar al siguiente. Utilice valores apropiados para su implementación de LDAP. *host*Sustitúyalo por la dirección IP o el nombre de host que se pueda resolver del servidor LDAP.

```
base dc=example,dc=org
uri ldap://host
ldap_version 3
binddn cn=admin,dc=example,dc=org
bindpw admin
```

Habilitar el complemento de autenticación de LDAP para JupyterHub

Utilice un editor de texto para modificar el archivo `/etc/jupyter/conf/jupyterhub_config.py` y añadir propiedades ldapauthenticator similares a las siguientes. Sustituir *host* por la dirección IP o el nombre de host resoluble del servidor LDAP. En el ejemplo se supone que los objetos de usuario se encuentran dentro de una unidad organizativa denominada *people* ou y se utilizan los componentes de nombres distintivos que se establecieron anteriormente. `ldap.conf`

```
c.JupyterHub.authenticator_class = 'ldapauthenticator.LDAPAuthenticator'
c.LDAPAuthenticator.use_ssl = False
c.LDAPAuthenticator.server_address = 'host'
c.LDAPAuthenticator.bind_dn_template = 'cn={username},ou=people,dc=example,dc=org'
```

Configuración de LDAP en el contenedor

Utilice un editor de texto para crear un script bash con los siguientes contenidos:

```
#!/bin/bash

# Uncomment the following lines to install LDAP client libraries only if
# using Amazon EMR release version 5.14.0. Later versions install libraries by default.
# sudo docker exec jupyterhub bash -c "sudo apt-get update"
# sudo docker exec jupyterhub bash -c "sudo apt-get -y install libnss-ldap libpam-ldap
 ldap-utils nscd"

# Copy ldap.conf
sudo docker cp ldap.conf jupyterhub:/etc/ldap/
sudo docker exec jupyterhub bash -c "cat /etc/ldap/ldap.conf"

# configure nss switch
sudo docker exec jupyterhub bash -c "sed -i 's/\(^passwd.*\)/\1 ldap/g' /etc/
nsswitch.conf"
```

```
sudo docker exec jupyterhub bash -c "sed -i 's/\(^group.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "sed -i 's/\(^shadow.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "cat /etc/nsswitch.conf"

# configure PAM to create home directories
sudo docker exec jupyterhub bash -c "echo 'session required        pam_mkhomedir.so
 skel=/etc/skel umask=077' >> /etc/pam.d/common-session"
sudo docker exec jupyterhub bash -c "cat /etc/pam.d/common-session"

# restart nscd service
sudo docker exec jupyterhub bash -c "sudo service nscd restart"

# Test
sudo docker exec jupyterhub bash -c "getent passwd"

# Install ldap plugin
sudo docker exec jupyterhub bash -c "pip install jupyterhub-ldapauthenticator"
```

Guarde el script en el nodo principal y, a continuación, ejecútelo desde la línea de comandos del nodo principal. Por ejemplo, con el script guardado como `configure_ldap_client.sh` haga que el archivo sea ejecutable:

```
chmod +x configure_ldap_client.sh
```

Y ejecute el script:

```
./configure_ldap_client.sh
```

Adición de atributos a Active Directory

Para encontrar a cada usuario y crear la entrada adecuada en la base de datos, el contenedor JupyterHub docker requiere las siguientes propiedades de UNIX para el objeto de usuario correspondiente en Active Directory. Para más información, consulte la sección ¿Cómo puedo continuar editando los atributos GID/UID RFC 2307 ahora que el complemento de atributos de Unix ya no está disponible para el complemento de MMC de computadoras y usuarios de Active Directory? en el artículo Aclaración en relación con el estado de Identity Management for Unix (IDMU) y el rol de NIS Server en Windows Server 2016 Technical Preview y más allá.

- `homeDirectory`

Esta es la ubicación del directorio de inicio del usuario, que suele ser /home/*username*.

- `gidNumber`

  Es un valor mayor que 60 000 que todavía no esté utilizando ningún otro usuario. Compruebe en el archivo `etc/passwd` cuáles son los gids en uso.

- `uidNumber`

  Es un valor mayor que 60 000 que todavía no esté utilizando ningún otro grupo. Compruebe en el archivo `etc/group` cuáles son los uids en uso.

- `uid`

  Es lo mismo que el. *username*

Crear directorios de inicio de los usuarios

JupyterHub necesita directorios de inicio en el contenedor para autenticar los usuarios LDAP y almacenar datos de instancia. El siguiente ejemplo muestra dos usuarios, shirley y diego, en el directorio LDAP.

El primer paso consiste en consultar al servidor LDAP la información sobre el identificador de usuario y el identificador de grupo de cada usuario mediante ldapsearch, como se muestra en el siguiente ejemplo, y sustituirlo por la dirección IP o el nombre de *host* host resoluble del servidor LDAP:

```
ldapsearch -x -H ldap://host \
 -D "cn=admin,dc=example,dc=org" \
 -w admin \
 -b "ou=people,dc=example,dc=org" \
 -s sub \
 "(objectclass=*)" uidNumber gidNumber
```

El comando `ldapsearch` devuelve una respuesta con formato LDIF que se parece a la siguiente para los usuarios shirley y diego.

```
# extended LDIF

# LDAPv3
# base <ou=people,dc=example,dc=org> with scope subtree
# filter: (objectclass=*)
```

```
# requesting: uidNumber gidNumber sn

# people, example.org
dn: ou=people,dc=example,dc=org

# diego, people, example.org
dn: cn=diego,ou=people,dc=example,dc=org
sn: B
uidNumber: 1001
gidNumber: 100

# shirley, people, example.org
dn: cn=shirley,ou=people,dc=example,dc=org
sn: A
uidNumber: 1002
gidNumber: 100

# search result
search: 2
result: 0 Success

# numResponses: 4
# numEntries: 3
```

Con la información de la respuesta, ejecute comandos en el contenedor para crear un directorio de inicio para cada nombre común de usuario (`cn`). Utilice `uidNumber` y `gidNumber` para fijar la propiedad del directorio de inicio de ese usuario. Los siguientes comandos de ejemplo hacen esto para el usuario. *`shirley`*

```
sudo docker container exec jupyterhub bash -c "mkdir /home/shirley"
sudo docker container exec jupyterhub bash -c "chown -R $uidNumber /home/shirley"
sudo docker container exec jupyterhub bash -c "sudo chgrp -R $gidNumber /home/shirley"
```

> **Note**
>
> El autenticador LDAP no JupyterHub admite la creación de usuarios locales. Para obtener más información, consulte Nota de la configuración del autenticador LDAP sobre la creación de usuarios locales.
>
> Para crear manualmente un usuario local, utilice el siguiente comando.

```
sudo docker exec jupyterhub bash -c "echo 'shirley:x:$uidNumber:$gidNumber::/
home/shirley:/bin/bash' >> /etc/passwd"
```

Reiniciar el JupyterHub contenedor

Ejecute los siguientes comandos para reiniciar el contenedor de `jupyterhub`:

```
sudo docker stop jupyterhub
sudo docker start jupyterhub
```

Suplantación de usuarios

Un trabajo de Spark que se ejecute dentro de un cuaderno de Jupyter recorre varias aplicaciones durante su ejecución en Amazon EMR. Por ejemplo, Sparkmagic recibe PySpark tres códigos que un usuario ejecuta en Jupyter, y utiliza una solicitud HTTP POST para enviarlos a Livy, que luego crea un trabajo de Spark para ejecutarlo en el clúster mediante YARN.

De forma predeterminada, los trabajos de YARN enviados de esta manera se ejecutan como el usuario `livy`, con independencia de quién haya iniciado el trabajo. Al configurar la suplantación de usuarios, puede hacer que el ID de usuario del usuario del bloc de notas también sea el usuario asociado al trabajo de YARN. En lugar de que los trabajos iniciados por `shirley` y `diego` se asocien al usuario `livy`, los trabajos que inicia cada usuario se asocian a `shirley` y `diego`, respectivamente. Esto le ayuda a auditar el uso de Jupyter y a administrar las aplicaciones dentro de su organización.

Esta configuración solo se admite cuando las llamadas de Sparkmagic a Livy no están autenticadas. Las aplicaciones que proporcionan una capa de autenticación o delegación entre las aplicaciones de Hadoop y Livy (como Apache Knox Gateway) no se admiten. En los pasos de esta sección para configurar la suplantación de identidad de un usuario, se parte del supuesto de que JupyterHub Livy y Livy se ejecutan en el mismo nodo maestro. Si la aplicación tiene clústeres distintos, Paso 3: Crear directorios de inicio de HDFS para los usuarios debe modificarse de manera que los directorios HDFS se creen en el nodo principal de Livy.

Pasos para configurar la suplantación de usuarios

- Paso 1: Configurar Livy
- Paso 2: Añadir usuarios

- Paso 3: Crear directorios de inicio de HDFS para los usuarios

Paso 1: Configurar Livy

Puede utilizar las clasificaciones de configuración `livy-conf` y `core-site` al crear un clúster para habilitar la suplantación de usuarios de Livy tal y como se muestra en el siguiente ejemplo. Guarde la clasificación de configuración como JSON y haga referencia a ella cuando cree el clúster o especifique la clasificación de configuración directamente. Para obtener más información, consulte Configuración de aplicaciones.

```
[
  {
    "Classification": "livy-conf",
    "Properties": {
      "livy.impersonation.enabled": "true"
    }
  },
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.proxyuser.livy.groups": "*",
      "hadoop.proxyuser.livy.hosts": "*"
    }
  }
]
```

Paso 2: Añadir usuarios

Agregue JupyterHub usuarios mediante PAM o LDAP. Para obtener más información, consulte Uso de la autenticación PAM y Uso de la autenticación LDAP.

Paso 3: Crear directorios de inicio de HDFS para los usuarios

Se ha conectado al nodo principal para crear usuarios. Mientras está conectado al nodo principal, copie el contenido que se indica a continuación y guárdelo en un archivo de script. El script crea directorios principales de HDFS para cada JupyterHub usuario del nodo principal. El script asume que está utilizando el ID de usuario administrador predeterminado,*jovyan*.

```
#!/bin/bash

CURL="curl --silent -k"
```

```
HOST=$(curl -s http://169.254.169.254/latest/meta-data/local-hostname)

admin_token() {
    local user=jovyan
    local pwd=jupyter
    local token=$($CURL https://$HOST:9443/hub/api/authorizations/token \
        -d "{\"username\":\"$user\", \"password\":\"$pwd\"}" | jq ".token")
    if [[ $token != null ]]; then
        token=$(echo $token | sed 's/"//g')
    else
        echo "Unable to get Jupyter API Token."
        exit 1
    fi
    echo $token
}

# Get Jupyter Admin token
token=$(admin_token)

# Get list of Jupyter users
users=$(curl -XGET -s -k https://$HOST:9443/hub/api/users \
 -H "Authorization: token $token" | jq '.[].name' | sed 's/"//g')

# Create HDFS home dir
for user in ${users[@]};
do
 echo "Create hdfs home dir for $user"
 hadoop fs -mkdir /user/$user
 hadoop fs -chmod 777 /user/$user
done
```

Instalación de kernels y bibliotecas adicionales

Al crear un clúster JupyterHub en Amazon EMR, el kernel de Python 3 predeterminado para Jupyter PySpark y los kernels para Spark para Sparkmagic se instalan en el contenedor de Docker. Puede instalar kernels adicionales. También puede instalar bibliotecas y paquetes adicionales y, a continuación, importarlos al shell apropiado.

Instalación de un kernel

Los kernels se instalan en el contenedor de Docker. La manera más fácil de conseguirlo es crear un script bash con comandos de instalación, guardarlo en el nodo principal y, a continuación,

usar el comando `sudo docker exec jupyterhub` *`script_name`* para ejecutar el script en el contenedor `jupyterhub`. El siguiente script de ejemplo instala el kernel y, a continuación, instala algunas bibliotecas para ese kernel en el nodo principal para que más tarde pueda importar las bibliotecas utilizando el kernel en Jupyter.

```
#!/bin/bash

# Install Python 2 kernel
conda create -n py27 python=2.7 anaconda
source /opt/conda/envs/py27/bin/activate
apt-get update
apt-get install -y gcc
/opt/conda/envs/py27/bin/python -m pip install --upgrade ipykernel
/opt/conda/envs/py27/bin/python -m ipykernel install

# Install libraries for Python 2
/opt/conda/envs/py27/bin/pip install paramiko nltk scipy numpy scikit-learn pandas
```

Para instalar el kernel y las bibliotecas en el contenedor, abra una conexión de terminal al nodo principal, guarde el script en `/etc/jupyter/install_kernels.sh` y ejecute el siguiente comando en la línea de comandos del nodo principal:

```
sudo docker exec jupyterhub bash /etc/jupyter/install_kernels.sh
```

Uso de bibliotecas e instalación de bibliotecas adicionales

En Amazon EMR se preinstala un conjunto principal de bibliotecas de machine learning y de ciencia de datos para Python 3. JupyterHub Puede utilizar `sudo docker exec jupyterhub bash -c "conda list"` y `sudo docker exec jupyterhub bash -c "pip freeze"`.

Si un trabajo de Spark necesita bibliotecas en los nodos de trabajo, le recomendamos que utilice una acción de arranque para ejecutar un script que instale las bibliotecas cuando cree el clúster. Las acciones de arranque se ejecutan en todos los nodos del clúster durante el proceso de creación de clúster, lo que simplifica la instalación. Si instala bibliotecas en nodos secundarios o de trabajo después un clúster esté en ejecución, la operación es más compleja. Proporcionamos un ejemplo programa en Python en esta sección que muestra cómo instalar estas bibliotecas.

La acción de arranque y los ejemplos de programas de Python mostrados en esta sección utilizan un script bash guardado en Amazon S3 para instalar las bibliotecas en todos los nodos.

El script al que se hace referencia en los siguientes ejemplos usa `pip` para instalar paramiko, nltk, scipy, scikit-learn y pandas para el kernel de Python 3:

```
#!/bin/bash

sudo python3 -m pip install boto3 paramiko nltk scipy scikit-learn pandas
```

Después de crear el script, cárguelo en una ubicación en Amazon S3, por ejemplo, `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`. Para obtener más información, consulte Carga de objetos en la Guía del usuario de Amazon Simple Storage Service para que pueda utilizarlos en su acción de arranque o en su programa de Python.

Para especificar una acción de arranque que instale bibliotecas en todos los nodos al crear un clúster con la AWS CLI

1. Cree un script similar al ejemplo anterior y guárdelo en una ubicación en Amazon S3. Usamos el ejemplo `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`.
2. Cree el clúster con JupyterHub el `Path` argumento de la `--bootstrap-actions` opción para especificar la ubicación del script como se muestra en el siguiente ejemplo:

   > **Note**
   >
   > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

   ```
   aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
   --applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/
   MyJupyterClusterLogs \
   --use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
    KeyName=MyKeyPair \
   --bootstrap-actions Path=s3://amzn-s3-demo-bucket/install-my-jupyter-
   libraries.sh,Name=InstallJupyterLibs
   ```

Para especificar una acción de arranque que instale bibliotecas en todos los nodos cuando cree un clúster mediante la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. Especifique la configuración de Software and Steps (Software y pasos) y de Hardware según sea apropiado para su aplicación.
4. En la pantalla General Cluster Settings (Configuración de clúster general), expanda Bootstrap Actions (Acciones de arranque).
5. En Add bootstrap action (Añadir acción de arranque), seleccione Custom action (Acción personalizada), Configure and add (Configurar y añadir).
6. En Name (Nombre), escriba un nombre fácil de recordar. En Ubicación del script, escriba la ubicación en Amazon S3 de su script (el ejemplo que usamos es s3://amzn-s3-demo-bucket/install-my-jupyter-libraries .sh). Deje Optional arguments (Argumentos opcionales) en blanco y elija Add (Añadir).
7. Especifique otros ajustes para el clúster y elija Next (Siguiente).
8. Especifique la configuración de seguridad y seleccione Create cluster (Crear clúster).

Example Instalación de bibliotecas en los nodos principales de un clúster en ejecución

Después de instalar las bibliotecas en el nodo principal desde Jupyter, puede instalar bibliotecas en nodos secundarios en ejecución de varias maneras. El siguiente ejemplo muestra un programa Python escrito para ejecutarse en un equipo local. Cuando ejecutas el programa Python de forma local, utiliza `AWS-RunShellScript` of AWS Systems Manager para ejecutar el script de ejemplo, que se muestra anteriormente en esta sección, que instala bibliotecas en los nodos principales del clúster.

```
import argparse
import time
import boto3


def install_libraries_on_core_nodes(cluster_id, script_path, emr_client, ssm_client):
    """
    Copies and runs a shell script on the core nodes in the cluster.
```

```
    :param cluster_id: The ID of the cluster.
    :param script_path: The path to the script, typically an Amazon S3 object URL.
    :param emr_client: The Boto3 Amazon EMR client.
    :param ssm_client: The Boto3 AWS Systems Manager client.
    """
    core_nodes = emr_client.list_instances(
        ClusterId=cluster_id, InstanceGroupTypes=["CORE"]
    )["Instances"]
    core_instance_ids = [node["Ec2InstanceId"] for node in core_nodes]
    print(f"Found core instances: {core_instance_ids}.")

    commands = [
        # Copy the shell script from Amazon S3 to each node instance.
        f"aws s3 cp {script_path} /home/hadoop",
        # Run the shell script to install libraries on each node instance.
        "bash /home/hadoop/install_libraries.sh",
    ]
    for command in commands:
        print(f"Sending '{command}' to core instances...")
        command_id = ssm_client.send_command(
            InstanceIds=core_instance_ids,
            DocumentName="AWS-RunShellScript",
            Parameters={"commands": [command]},
            TimeoutSeconds=3600,
        )["Command"]["CommandId"]
        while True:
            # Verify the previous step succeeded before running the next step.
            cmd_result = ssm_client.list_commands(CommandId=command_id)["Commands"][0]
            if cmd_result["StatusDetails"] == "Success":
                print(f"Command succeeded.")
                break
            elif cmd_result["StatusDetails"] in ["Pending", "InProgress"]:
                print(f"Command status is {cmd_result['StatusDetails']}, waiting...")
                time.sleep(10)
            else:
                print(f"Command status is {cmd_result['StatusDetails']}, quitting.")
                raise RuntimeError(
                    f"Command {command} failed to run. "
                    f"Details: {cmd_result['StatusDetails']}"
                )


def main():
```

```
    parser = argparse.ArgumentParser()
    parser.add_argument("cluster_id", help="The ID of the cluster.")
    parser.add_argument("script_path", help="The path to the script in Amazon S3.")
    args = parser.parse_args()

    emr_client = boto3.client("emr")
    ssm_client = boto3.client("ssm")

    install_libraries_on_core_nodes(
        args.cluster_id, args.script_path, emr_client, ssm_client
    )


if __name__ == "__main__":
    main()
```

JupyterHub historial de lanzamientos

En la tabla siguiente, se muestra la versión de JupyterHub incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

JupyterHub información sobre la versión

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr7.9.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server,

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
		spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr7.8.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.7.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-7.6.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.5.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-7.4.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.3.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-7.2.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.2	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-7.1.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.0.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.15.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.14.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.13.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.12.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.11.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.11.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.10.1	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.10.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.9.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.9.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.8.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.8.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.7.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.36.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.6.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.35.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.5.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.4.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.3.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.3.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.2.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.2.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.1.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.1.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.0.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-6.0.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.34.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.33.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.33.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.32.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.32.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.31.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.31.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.30.2	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.30.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.30.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.29.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.28.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.28.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.27.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.27.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.26.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.25.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.24.1	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.24.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.23.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.23.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.22.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.2	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.21.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.20.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.20.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.19.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.19.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.18.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.18.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.17.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.17.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.17.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.16.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.16.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.15.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.15.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Etiqueta de versión de Amazon EMR	JupyterHub Versión	Componentes instalados con JupyterHub
emr-5.14.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Apache Livy

Livy permite la interacción a través de una interfaz de REST con un clúster de EMR en el que se ejecuta Spark. Puede utilizar la interfaz de REST o una biblioteca de cliente RPC para enviar trabajos de Spark o fragmentos de código de Spark, recuperar resultados de forma síncrona o asíncrona, y gestionar Spark Context. Para obtener más información, consulte el sitio web de Apache Livy. Livy se incluye en la versión de lanzamiento 5.9.0 de Amazon EMR y posteriores.

Para acceder a la interfaz web de Livy, configure un túnel de SSH en el nodo maestro y una conexión proxy. Para obtener más información, consulte Ver interfaces web alojadas en clústeres de EMR.

En la siguiente tabla, se muestra la versión de Livy incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Livy.

Para ver la versión de los componentes instalados con Livy en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Livy para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-7.9.0	Livy 0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

En la tabla siguiente, se muestra la versión de Livy incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Livy.

Para ver la versión de los componentes instalados con Livy en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Livy para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.15.0	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

En la tabla siguiente, se muestra la versión de Livy incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Livy.

Para ver la versión de los componentes instalados con Livy en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Livy para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.36.2	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
		yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Temas

- Habilitación de HTTPS con Apache Livy
- Historial de versiones de Livy

Habilitación de HTTPS con Apache Livy

Este tema es relevante si está ejecutando Amazon EMR 7.3.0 o una versión anterior. A partir de la versión 7.4.0, HTTPS está habilitado con Apache Livy de forma predeterminada.

1. Aprovisione un clúster de Amazon EMR con el cifrado de tránsito activado. Para obtener más información sobre el cifrado, consulte Cifrado de datos en reposo y en tránsito.
2. Cree un archivo denominado `livy_ssl.sh` con el siguiente contenido.

```
#!/bin/bash

KEYSTORE_FILE=`awk '/ssl.server.keystore.location/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEYSTORE_PASS=`awk '/ssl.server.keystore.password/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEY_PASS=`awk '/ssl.server.keystore.keypassword/{getline; print}' /etc/hadoop/conf/
ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`

echo "livy.keystore $KEYSTORE_FILE
livy.keystore.password $KEYSTORE_PASS
livy.key-password $KEY_PASS" | sudo tee -a /etc/livy/conf/livy.conf >/dev/null
```

```
sudo systemctl restart livy-server.service
```

3. Ejecute el siguiente script como un paso de Amazon EMR. Este script modifica `/etc/livy/conf/livy.conf` para activar SSL.

```
--steps '[{"Args":["s3://amzn-s3-demo-bucket/
livy_ssl.sh"],"Type":"CUSTOM_JAR","ActionOnFailure":"CONTINUE","Jar":"s3://
us-east-1.elasticmapreduce/libs/script-runner/script-
runner.jar","Properties":"","Name":"Custom JAR"}]'
```

4. Vuelva a iniciar el servicio Apache Livy para que los cambios se apliquen. Para reiniciar Apache Livy, consulte Detener y reiniciar procesos.
5. Compruebe que los clientes ya puedan comunicarse mediante HTTPS. Para enviar un trabajo, por ejemplo, ejecute el siguiente código.

```
curl -k -X POST --data '{"file": "local:///usr/lib/spark/examples/jars/spark-
examples.jar",
"className": "org.apache.spark.examples.SparkPi"}' \
-H "Content-Type: application/json" \
https://EMR_Master_Node_Host:8998/batches
```

Si ha activado HTTPS correctamente, Livy envía una respuesta en la que se indica que el comando se ha aceptado y que se ha enviado el trabajo por lotes.

```
{"id":1,"name":null,"owner":null,"proxyUser":null,"state":"starting","appId":null,"appInfo"
{"driverLogUrl":null,"sparkUiUrl":null},"log":["stdout: ","\nstderr: ","\nYARN
 Diagnostics: "]}
```

Historial de versiones de Livy

En la siguiente tabla, se muestra la versión de Livy incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Livy

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-7.9.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.8.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.7.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager,

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
		hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.6.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.5.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-7.4.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.3.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-7.2.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.2	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-7.1.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.0.0	0.7.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.15.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.14.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.13.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.12.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.11.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.11.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.10.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.10.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.9.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.9.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.8.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.8.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.7.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.36.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.6.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.35.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.5.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.4.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.3.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.3.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.2.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.2.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.1.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.1.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.0.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-6.0.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.34.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.33.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.33.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.32.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.32.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.31.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.31.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.30.2	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.30.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.30.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.29.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.28.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.28.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.27.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.27.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.26.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.25.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.24.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.24.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.23.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.23.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.22.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.21.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.20.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.20.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.19.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.19.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.18.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.18.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.17.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.17.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.17.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.16.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.16.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.15.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.15.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.14.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.13.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.13.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.12.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.12.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.11.4	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.11.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.11.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.10.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.10.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.9.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Etiqueta de versión de Amazon EMR	Versión de Livy	Componentes instalados con Livy
emr-5.9.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Apache MXNet

> **Note**
>
> MXNet ha llegado al final de su vida útil y el proyecto está archivado. Amazon EMR 7.2.0 es la última versión que incluye. MXNet

Apache MXNet es una biblioteca de aceleración diseñada para crear redes neuronales y otras aplicaciones de aprendizaje profundo. MXNetautomatiza los flujos de trabajo comunes y optimiza los cálculos numéricos. MXNet le ayuda a diseñar arquitecturas de redes neuronales sin tener que centrarse en implementar cálculos de bajo nivel, como las operaciones de álgebra lineal. MXNet se incluye en la versión 5.10.0 y posteriores de Amazon EMR.

Para obtener más información, consulte el sitio web de Apache MXNet .

En la siguiente tabla se muestra la versión MXNet incluida en la última versión de la serie Amazon EMR 7.x, junto con los componentes con los que se instala Amazon EMR. MXNet

Para ver la versión de los componentes con los que se instala MXNet, consulte las versiones de componentes de la versión 7.2.0.

MXNet información sobre la versión de emr-7.2.0

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-7.2.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

En la siguiente tabla se muestra la versión MXNet incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes con los que se instala Amazon EMR. MXNet

Para ver la versión de los componentes que se incluyen MXNet en esta versión, consulte Versiones de componentes de la versión 6.15.0.

MXNet información sobre la versión de emr-6.15.0

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.15.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

En la siguiente tabla se muestra la versión MXNet incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes con los que se instala Amazon EMR. MXNet

Para ver la versión de los componentes que se incluyen MXNet en esta versión, consulte Versiones de componentes de la versión 5.36.2.

MXNet información sobre la versión de emr-5.36.2

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.36.2	MXNet 1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
		cemanager, hadoop-yarn-timeline-server, mxnet, opencv

MXNet historial de versiones

En la siguiente tabla se muestra la versión MXNet incluida en cada versión de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

MXNet información sobre la versión

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-7.2.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.2	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
		cemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.1.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.0.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.15.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.14.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.13.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.12.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.11.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.11.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.10.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.10.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.8.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.8.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.7.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.36.1	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.6.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.35.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.5.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.4.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.3.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.3.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.2.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.2.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-6.0.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.0.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.34.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.33.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.33.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.32.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.32.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.30.2	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.29.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.27.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.27.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.26.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.25.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.23.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.23.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.22.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.21.2	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.20.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.20.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.19.1	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.19.0	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.17.2	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.16.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.16.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.15.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.15.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.2	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.14.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.13.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.13.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.12.3	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.2	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.12.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.4	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.3	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.11.2	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Etiqueta de versión de Amazon EMR	MXNet Versión	Componentes instalados con MXNet
emr-5.10.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.10.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Apache Oozie

> **Note**
>
> Obsolescencia de Oozie: dado que OSS Oozie ha pasado al modo ático, se eliminará de una de las próximas versiones de EMR. Recomendamos a los clientes que utilicen MWAA como alternativa a Oozie. Puede consultar la utilidad para migrar los flujos de trabajo oozie-to-airflow-emrde Oozie existentes GitHub a la MWAA.

Utilice el Programador de flujo de trabajo de Apache Oozie para administrar y coordinar los trabajos de Hadoop. Para obtener más información, consulte https://attic.apache.org/projects/oozie.html.

La interfaz web nativa de Oozie no es compatible con Amazon EMR. Para utilizar una interfaz front-end para Oozie, pruebe la aplicación Oozie Hue. Para obtener más información, consulte Hue. Oozie se incluye en la versión de lanzamiento de Amazon EMR 5.0.0 y posteriores. Oozie se incluye como una aplicación de entorno de pruebas en las versiones anteriores. Para obtener más información, consulte Versiones de lanzamiento de Amazon EMR 4.x.

Si utiliza una AMI personalizada de Amazon Linux basada en una AMI de Amazon Linux con una fecha de creación del 11 de agosto de 2018, el servidor de Oozie no podrá iniciarse. Si utiliza Oozie, cree una AMI personalizada basada en un ID de AMI de Amazon Linux con una fecha de creación diferente. Puede utilizar el siguiente AWS CLI comando de la para obtener una lista de imágenes IDs de todas las HVM de Amazon Linux AMIs con la versión 2018.03, junto con la fecha de lanzamiento, de modo que pueda elegir una AMI de Amazon Linux adecuada como base. Sustituya por MyRegion el identificador de su región, como us-west-2.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

En la siguiente tabla, se muestra la versión de Oozie incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Oozie.

Para ver la versión de los componentes instalados con Oozie en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Oozie para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-7.9.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

En la tabla siguiente, se muestra la versión de Oozie incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Oozie.

Para ver la versión de los componentes instalados con Oozie en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Oozie para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.15.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
		nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

En la tabla siguiente, se muestra la versión de Oozie incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Oozie.

Para ver la versión de los componentes instalados con Oozie en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Oozie para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.36.2	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Temas

- Uso de Oozie con una base de datos remota en Amazon RDS
- Configurar la versión Java para Oozie

- Historial de versiones de Oozie

Uso de Oozie con una base de datos remota en Amazon RDS

De forma predeterminada, la información de usuario de Oozie y las historias de consulta se almacenan en una base de datos MySQL local en el nodo principal. También puede crear uno o varios clústeres habilitados para Oozie con una configuración almacenada en Amazon S3 y una base de datos MySQL en Amazon Relational Database Service (Amazon RDS). Esto le permite conservar información de usuario y el historial de consultas creado por Oozie sin mantener el clúster de Amazon EMR en ejecución. Le recomendamos utilizar el cifrado del servidor de Amazon S3 para almacenar el archivo de configuración.

En primer lugar, cree la base de datos remota para Oozie.

Para crear una base de datos MySQL externa

1. Abra la consola de Amazon RDS en https://console.aws.amazon.com/rds/.
2. Seleccione Launch a DB Instance (Lanzar una instancia de base de datos).
3. Elija MySQL y, a continuación, elija Select (Seleccionar).
4. Deje la selección predeterminada de Multi-AZ Deployment and Provisioned IOPS Storage (Implementación Multi-AZ y almacenamiento de IOPS provisionadas) y elija Next (Siguiente).
5. Deje los valores predeterminados de Instance Specifications (Especificaciones de la instancia), especifique Settings (Configuración) y elija Next (Siguiente).
6. En la página Configure Advanced Settings (Configurar opciones avanzadas), elija un grupo de seguridad y nombres de bases de datos adecuados. El grupo de seguridad que utilice debe, al menos, permitir el acceso TCP de entrada al puerto 3306 desde el nodo principal del clúster. Si no ha creado su clúster en este punto, puede permitir que todos los hosts se conecten al puerto 3306 y ajustar el grupo de seguridad después de haber lanzado el clúster. Elija Launch DB Instance (Lanzar instancia de base de datos).
7. En el panel de RDS, seleccione Instances (Instancias) y seleccione la instancia que acaba de crear. Cuando la base de datos esté disponible, anote el dbname, nombre de usuario, contraseña y nombre de host de la instancia de RDS. Utilizará esta información al crear y configurar el clúster.

Para especificar una base de datos MySQL externa para Oozie al lanzar un clúster mediante AWS CLI

Para especificar una base de datos de MySQL externa para Oozie al lanzar un clúster con la AWS CLI, utilice la información que ha anotado al crear la instancia de RDS para configurar `oozie-site` con un objeto de configuración.

> **Note**
>
> Puede crear varios clústeres que utilicen la misma base de datos externa, pero cada clúster compartirá el historial de consultas y la información de usuario.

- Utilice el AWS CLI, cree un clúster con Oozie instalado, utilizando la base de datos externa que creó y haciendo referencia a un archivo de configuración con una clasificación de configuración para Oozie que especifique las propiedades de la base de datos. En el siguiente ejemplo, se crea un clúster con Oozie instalado, que hace referencia a un archivo de configuración de Amazon S3, `myConfig.json`, que especifica la configuración de la base de datos.

  > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Oozie
   Name=Spark Name=Hive \
  --instance-type m5.xlarge --instance-count 3 \
  --configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
  myConfig.json --use-default-roles
  ```

  A continuación, se muestra contenido de ejemplo del archivo `myConfig.json`. Sustituya y *`password`* por la URL *`JDBC URL`*`username`, el nombre de usuario y la contraseña de JDBC de su instancia de RDS.

> **Important**
>
> La dirección URL de JDBC debe incluir el nombre de la base de datos como sufijo. Por ejemplo, jdbc:mysql://.xxxxxxxxxx.us-east-1.rds.amazonaws.com:3306/ oozie-external-db dbname.

```
[{
  "Classification": "oozie-site",
    "Properties": {
        "oozie.service.JPAService.jdbc.driver": "org.mariadb.jdbc.Driver",
        "oozie.service.JPAService.jdbc.url": "JDBC URL",

        "oozie.service.JPAService.jdbc.username": "username",
        "oozie.service.JPAService.jdbc.password": "password"
    },
    "Configurations": []
}]
```

Configurar la versión Java para Oozie

Oozie ejecuta varios procesos de máquina virtual Java (JVM). En esta página, se explica cómo configurar la versión de Java para cada proceso.

- Servidor de Oozie: se establece `JAVA_HOME` en la clasificación `oozie-env` para actualizar la versión de Java para `EmbeddedOozieServer`.
- Oozie Launcher AM: Oozie Launcher AM es un trabajo de MR con un solo asignador que invoca las bibliotecas cliente de aplicaciones adecuadas, como Hadoop y Hive. A menos que se configure lo contrario, las versiones de tiempo de ejecución de Oozie Launcher AM son las mismas que las de los tiempos de ejecución de Java para Hadoop en el clúster de EMR. Para configurar el tiempo de ejecución de Java para Oozie Launcher AM, defina la siguiente propiedad en el `workflow.xml` para el trabajo:

```
<property>
     <name>mapred.child.env</name>
     <value>JAVA_HOME=/path/to/JAVA_HOME</value>
 </property>
```

Esta propiedad garantiza que Oozie Launcher AM para la tarea de Oozie se ejecute en la versión de Java que especifique, en lugar de en la versión de Java configurada en Hadoop.

- Ejecutable en el cliente de la aplicación: dado que Oozie Launcher AM invoca el cliente de la aplicación de forma predeterminada, el tiempo de ejecución de Java del ejecutable del cliente es el mismo que el de Oozie Launcher AM.
- Aplicaciones lanzadas por un trabajo de Oozie: a menos que se especifique lo contrario, las versiones de tiempo de ejecución de la aplicación reales JVMs que se lanzan mediante un trabajo de Oozie son las mismas que las de los tiempos de ejecución de Java para Hadoop en el clúster EMR. Según el tipo de acción de flujo de trabajo de Oozie que se utilice para iniciar la aplicación en un trabajo de Oozie (acción de Spark o Hive), puede actualizar el tiempo de ejecución de Java predeterminado para la aplicación real JVMs en el `workflow.xml` trabajo de Oozie.

Historial de versiones de Oozie

En la siguiente tabla, se muestra la versión de Oozie incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Oozie

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client,

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
		oozie-server, tez-on-yarn, tez-on-worker
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-7.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-7.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-7.2.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-5.36.2	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-7.1.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.0.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.15.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.14.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.13.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.12.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.11.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.11.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.10.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.10.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.9.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.8.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.36.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.36.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.35.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.3.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.2.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.2.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.1.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.1.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.0.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-6.0.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.34.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.33.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.33.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.32.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.32.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.31.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.31.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.30.2	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.30.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.30.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.29.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.28.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.28.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.27.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.27.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.26.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.25.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.24.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.24.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.23.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.23.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.22.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.21.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.20.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.20.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.19.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.19.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.18.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.18.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.17.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.17.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.17.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.16.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.16.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.15.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.15.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.14.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.13.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.13.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.12.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.12.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.11.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.11.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.11.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.10.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.10.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.9.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.9.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.8.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.8.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.7.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.7.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.6.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.6.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.5.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.5.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.4.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.4.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.3.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.3.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.3.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.2.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.1.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.1.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.0.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Oozie	Componentes instalados con Oozie
emr-5.0.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Notas de la versión de Oozie por versión

Temas

- Amazon EMR 7.6.0: cambios en Oozie
- Amazon EMR 7.5.0: cambios en Oozie
- Amazon EMR 7.4.0: cambios en Oozie
- Amazon EMR 7.3.0: cambios en Oozie
- Amazon EMR 7.1.0: cambios en Oozie
- Amazon EMR 7.0.0: cambios en Oozie

Amazon EMR 7.6.0: cambios en Oozie

Sin cambios en la versión 7.6.0.

Amazon EMR 7.5.0: cambios en Oozie

Tipo	Descripción
Upgrade	Elimine la compatibilidad con la organización de trabajos de Sqoop, ya que Sqoop no está disponible en las nuevas versiones de EMR

Amazon EMR 7.4.0: cambios en Oozie

Cambios en Oozie: SIN CAMBIOS

Amazon EMR 7.3.0: cambios en Oozie

Tipo	Descripción
Upgrade	Actualiza la guía a la versión 4.0.
Cambio de comportamiento predeterminado	Define dos propiedades en `oozie-site.xml` : • `oozie.service.SparkConfigurationService.spark.configurations=*=/etc/spark/conf/` • `oozie.service.SparkConfigurationService.spark.configurations.blacklist=spark.yarn.dist.files` Estas propiedades permiten a Oozie elegir las configuraciones predeterminadas de Amazon EMR Spark al instalar Oozie junto con Spark.

Amazon EMR 7.1.0: cambios en Oozie

Tipo	Descripción
Upgrade	Actualiza Netty a la versión 4.1.100.Final.
Upgrade	Actualización de Jetty a 9.4.53.v20231009.

Amazon EMR 7.0.0: cambios en Oozie

Tipo	Descripción
Upgrade	El cliente y el servidor de Oozie ahora usan Java 17 de forma predeterminada. Consulte la Guía de versión de EMR 7.0.0 para obtener más información.

Apache Phoenix

Apache Phoenix se usa para OLTP y análisis operativo, lo que le permite usar consultas SQL estándar y JDBC APIs para trabajar con un almacén de respaldo de Apache HBase . Para obtener más información, consulte Phoenix in 15 minutes or less. Phoenix se incluye en la versión de lanzamiento de Amazon EMR 4.7.0 y posteriores.

Si actualiza desde una versión anterior de Amazon EMR a una versión 5.4.0 o posterior de Amazon EMR y utiliza la indexación secundaria, actualice los índices locales como se describe en la documentación de Apache Phoenix. Amazon EMR elimina las configuraciones requeridas de la clasificación de `hbase-site`, pero los índices se tienen que volver a rellenar. Se admite la actualización de índices en línea y fuera de línea. Las actualizaciones en línea son la opción predeterminada, lo que significa que los índices se rellenan al inicializar desde los clientes de Phoenix versión 4.8.0 o superior. Para especificar actualizaciones fuera de línea, defina la configuración `phoenix.client.localIndexUpgrade` en false en la clasificación `phoenix-site` y, a continuación, SSH en el nodo principal para ejecutar `psql [zookeeper] -1`.

En la siguiente tabla, se muestra la versión de Phoenix incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Phoenix.

Para ver la versión de los componentes instalados con Phoenix en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Phoenix para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.9.0	Phoenix 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
		timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-queryserver, zookeeper-client, zookeeper-server

En la tabla siguiente, se muestra la versión de Phoenix incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Phoenix.

Para ver la versión de los componentes instalados con Phoenix en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Phoenix para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.15.0	Phoenix 5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
		server, zookeeper-client, zookeeper-server

En la tabla siguiente, se muestra la versión de Phoenix incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Phoenix.

Para ver la versión de los componentes instalados con Phoenix en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Phoenix para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.36.2	Phoenix 4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Temas

- Creación de un clúster con Phoenix
- Clientes de Phoenix

- Historial de versiones de Phoenix

Creación de un clúster con Phoenix

Puede instalar Phoenix eligiendo la aplicación al crear un clúster en la consola o con la AWS CLI. En los siguientes procedimientos y ejemplos se muestra cómo crear un clúster con Phoenix y HBase. Para obtener más información sobre la creación de clústeres con la consola, incluidas Opciones avanzadas, consulte Planificar y configurar clústeres en la Guía de administración de Amazon EMR.

Para lanzar un clúster con Phoenix instalado utilizando las Quick Options (Opciones rápidas) para crear un clúster en la consola

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Create cluster (Crear clúster) para utilizar Quick Create (Creación rápida).
3. Para Software Configuration (Configuración de software), elija la versión más reciente adecuada para su aplicación. Phoenix aparece como opción solo cuando se ha seleccionado Amazon Release Version emr-4.7.0 (Versión de Amazon emr-4.7.0) o posterior.
4. Para las aplicaciones, elija la segunda opción HBase: HBase ***ver*** con Ganglia, Hadoop***ver***, Hive, Hue***ver***, Phoenix ***ver*** y. *ver ver* ZooKeeper *ver*
5. Seleccione las demás opciones que necesite y, a continuación, elija Create cluster (Crear clúster).

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

El siguiente ejemplo lanza un clúster con Phoenix instalado utilizando las opciones de configuración predeterminadas.

Para lanzar un clúster con Phoenix y utilizar el HBase AWS CLI

- Cree el clúster con el siguiente comando:

```
aws emr create-cluster --name "Cluster with Phoenix" --release-label emr-7.9.0 \
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Personalización de configuraciones de Phoenix

Al crear un clúster, configure Phoenix definiendo valores en `hbase-site.xml` utilizando la clasificación de configuración `hbase-site`.

Para obtener más información, consulte Configuración y ajuste en la documentación de Phoenix.

En el siguiente ejemplo, se ilustra cómo utilizar un archivo JSON almacenado en Amazon S3 para especificar el valor `false` para la propiedad `phoenix.schema.dropMetaData`. Se pueden especificar múltiples propiedades para una única clasificación. Para obtener más información, consulte Configuración de aplicaciones. El comando `create-cluster`, a continuación, hace referencia al archivo JSON como parámetro `--configurations`.

El contenido del archivo JSON guardado to /amzn-s3-demo-bucket/myfolder/myconfig .json es el siguiente.

```
[
    {
      "Classification": "hbase-site",
      "Properties": {
        "phoenix.schema.dropMetaData": "false"
      }
    }
  ]
```

El comando `create cluster` que hace referencia al archivo JSON se muestra en el siguiente ejemplo.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Phoenix \
Name=HBase --instance-type m5.xlarge --instance-count 2 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myconfig.json
```

> **Note**
>
> La solicitud de reconfiguración para cualquier clasificación de configuración de Phoenix solo se admite en la versión 5.23.0 y versiones posteriores de Amazon EMR, y no es compatible con la versión 5.21.0 o 5.22.0 de Amazon EMR. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución

Clientes de Phoenix

Puede conectarse a Phoenix utilizando un cliente JDBC creado con dependencias completas o utilizando el "cliente ligero" que utiliza el Phoenix Query Server y que solo se puede ejecutar en un nodo principal de un clúster (por ejemplo, mediante la utilización de un cliente SQL, un paso, línea de comando, enrutamiento de puertos SSH, etc.). Si se utiliza el cliente JDBC «robusto», este aún necesita tener acceso a todos los nodos del clúster, ya que se conecta directamente a HBase los servicios. El cliente de Phoenix "ligero" solo necesita acceso a Phoenix Query Server en un puerto predeterminado 8765. Existen varios scripts dentro de Phoenix que utilizan estos clientes.

Utilizar un paso de Amazon EMR para consultar con Phoenix

El siguiente procedimiento restaura una instantánea de los datos HBase y los utiliza para ejecutar una consulta de Phoenix. Puede ampliar este ejemplo o crear un nuevo script que aproveche los clientes de Phoenix para adaptarse a sus necesidades.

1. Cree un clúster con Phoenix instalado, utilizando el comando siguiente:

```
aws emr create-cluster --name "Cluster with Phoenix" --log-uri s3://amzn-s3-demo-
bucket/myLogFolder --release-label emr-7.9.0 \
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

2. Cree y, a continuación, cargue los siguientes archivos en Amazon S3:

   copySnapshot.sh

```
sudo su hbase -s /bin/sh -c 'hbase snapshot export \
 -D hbase.rootdir=s3://us-east-1.elasticmapreduce.samples/hbase-demo-customer-data/
snapshot/ \
-snapshot customer_snapshot1 \
```

```
-copy-to hdfs://masterDNSName:8020/user/hbase \
-mappers 2 -chuser hbase -chmod 700'
```

runQuery.sh

```
aws s3 cp s3://amzn-s3-demo-bucket/phoenixQuery.sql /home/hadoop/
/usr/lib/phoenix/bin/sqlline-thin.py http://localhost:8765 /home/hadoop/
phoenixQuery.sql
```

phoenixQuery.sql

> **Note**
>
> Solo debe incluir `COLUMN_ENCODED_BYTES=0` en el siguiente ejemplo si utiliza las versiones 5.26.0 y superiores de Amazon EMR.

```
CREATE VIEW "customer" (
pk VARCHAR PRIMARY KEY,
"address"."state" VARCHAR,
"address"."street" VARCHAR,
"address"."city" VARCHAR,
"address"."zip" VARCHAR,
"cc"."number" VARCHAR,
"cc"."expire" VARCHAR,
"cc"."type" VARCHAR,
"contact"."phone" VARCHAR)
COLUMN_ENCODED_BYTES=0;

CREATE INDEX my_index ON "customer" ("customer"."state") INCLUDE("PK",
 "customer"."city", "customer"."expire", "customer"."type");

SELECT "customer"."type" AS credit_card_type, count(*) AS num_customers FROM
 "customer" WHERE "customer"."state" = 'CA' GROUP BY "customer"."type";
```

Utilice AWS CLI para enviar archivos al bucket de S3:

```
aws s3 cp copySnapshot.sh s3://amzn-s3-demo-bucket/
aws s3 cp runQuery.sh s3://amzn-s3-demo-bucket/
aws s3 cp phoenixQuery.sql s3://amzn-s3-demo-bucket/
```

3. Cree una tabla con el siguiente paso enviado al clúster que creó en el paso 1:

   createTable.json

```
[
  {
    "Name": "Create HBase Table",
    "Args": ["bash", "-c", "echo $'create \"customer\",\"address\",\"cc\",\"contact
\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./createTable.json
```

4. Utilice `script-runner.jar` para ejecutar el script `copySnapshot.sh` que cargó anteriormente en su bucket de S3:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="HBase Copy Snapshot",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/copySnapshot.sh"]
```

   Esto ejecuta un MapReduce trabajo para copiar los datos de la instantánea al HDFS del clúster.

5. Restaure la instantánea que ha copiado en el clúster utilizando el paso siguiente:

   restoreSnapshot.json

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo $'disable \"customer\"; restore_snapshot
 \"customer_snapshot1\"; enable \"customer\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./restoreSnapshot.json
```

6. Utilice `script-runner.jar` para ejecutar el script `runQuery.sh` que cargó anteriormente en su bucket de S3:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Phoenix Run Query",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/runQuery.sh"]
```

   La consulta se ejecuta y devuelve los resultados al `stdout` del paso. Es posible que este paso tarde unos minutos en completarse.

7. Revise los resultados del `stdout` del paso en el URI de registro que utilizó al crear el clúster en el paso 1. El resultado debe ser similar a lo siguiente:

```
+-----------------------------------------+-----------------------------------+
|             CREDIT_CARD_TYPE            |              NUM_CUSTOMERS        |
+-----------------------------------------+-----------------------------------+
| american_express                        | 5728                              |
| dankort                                 | 5782                              |
| diners_club                             | 5795                              |
| discover                                | 5715                              |
| forbrugsforeningen                      | 5691                              |
| jcb                                     | 5762                              |
| laser                                   | 5769                              |
| maestro                                 | 5816                              |
| mastercard                              | 5697                              |
| solo                                    | 5586                              |
| switch                                  | 5781                              |
| visa                                    | 5659                              |
+-----------------------------------------+-----------------------------------+
```

Historial de versiones de Phoenix

En la siguiente tabla, se muestra la versión de Phoenix incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión

correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Phoenix

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-regionserver, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-queryserver, zookeeper-client, zookeeper-server
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-reg

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
		ion-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server
emr-7.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.6.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.5.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.4.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.3.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.2.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.36.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.1.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-7.0.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.15.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.14.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.13.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.12.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.11.1	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.11.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.10.1	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.10.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.9.1	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.9.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.8.1	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.8.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.7.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.36.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.36.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.6.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.35.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.5.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.4.0	51.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.3.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.3.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.2.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.2.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.1.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.1.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.0.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-6.0.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.34.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.33.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.33.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.32.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.32.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.31.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.31.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.30.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.30.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.30.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.29.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.28.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.28.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.27.1	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.27.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.26.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.25.0	4,14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.24.1	4,14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.24.0	4,14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.23.1	4,14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.23.0	4,14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.22.0	4,14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.21.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.20.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.20.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.19.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.19.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.18.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.18.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.17.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.17.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.17.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.16.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.16.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.15.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.15.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.14.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.13.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.13.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.12.3	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.12.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.11.4	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.11.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.10.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.10.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.9.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.8.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.7.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.6.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.6.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.4	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.5.3	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.2	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.5.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.4.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.3.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.3.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.3.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.2.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.2.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.1.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.1.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.0.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-5.0.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.6	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.9.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.9.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.9.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.8.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.8.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.8.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.7.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	Versión de Phoenix	Componentes instalados con Phoenix
emr-4.7.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Apache Pig

Apache Pig es una biblioteca de Apache de código abierto que se ejecuta sobre Hadoop y proporciona un lenguaje de scripting que puede utilizar para transformar grandes conjuntos de datos complejos sin tener que escribir código complejo en un lenguaje informático de nivel inferior, como Java. La biblioteca toma comandos tipo SQL escritos en un lenguaje llamado Pig Latin y los convierte en tareas de Tez basadas en gráficos acíclicos dirigidos (DAGs) o programas. MapReduce Pig funciona con datos estructurados y no estructurados en una variedad de formatos. Para obtener más información acerca de Pig, consulte http://pig.apache.org/.

Puede ejecutar los comandos de Pig de forma interactiva o en el modo por lotes. Para utilizar Pig de forma interactiva, cree una conexión SSH al nodo principal y envíe los comandos utilizando el shell de Grunt. Para utilizar Pig en modo por lotes, escriba sus scripts de Pig, cárguelos en Amazon S3 y envíelos como pasos de clúster. Para obtener más información sobre cómo enviar trabajo a un clúster, consulte Enviar un trabajo a un clúster en la Guía de administración de Amazon EMR.

Cuando utilice Pig para escribir el resultado en una HCatalog tabla en Amazon S3, deshabilite la escritura directa de Amazon EMR; para ello, establezca las `mapred.output.direct.EmrFileSystem` propiedades `mapred.output.direct.NativeS3FileSystem` y en. `false` Para obtener más información, consulte Usando HCatalog. En un script de Pig, puede utilizar los comandos `SET mapred.output.direct.NativeS3FileSystem false` y `SET mapred.output.direct.EmrFileSystem false`.

En la siguiente tabla, se muestra la versión de Pig incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Pig.

Para ver la versión de los componentes instalados con Pig en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Pig para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-7.9.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-ma

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
		pred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

En la tabla siguiente, se muestra la versión de Pig incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Pig.

Para ver la versión de los componentes instalados con Pig en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Pig para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.15.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

En la tabla siguiente, se muestra la versión de Pig incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Pig.

Para ver la versión de los componentes instalados con Pig en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Pig para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.36.2	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Temas

- Enviar trabajo de Pig
- Llamada a funciones definidas por el usuario desde Pig
- Historial de versiones de Pig

Enviar trabajo de Pig

En esta sección, se muestra el envío de un trabajo de Pig a un clúster de Amazon EMR. En los ejemplos siguientes se genera un informe que contiene el total de bytes transferidos, una lista de las 50 direcciones IP principales, una lista de los 50 remitentes externos principales y los 50 términos de búsqueda principales con Bing y Google. El script de Pig se encuentra en el bucket de Amazon S3 `s3://elasticmapreduce/samples/pig-apache/do-reports2.pig`. Los datos de entrada se

encuentran en el bucket `s3://elasticmapreduce/samples/pig-apache/input` de Amazon S3. La salida se guarda en un bucket de Amazon S3.

Enviar trabajo de Pig mediante la consola de Amazon EMR

En este ejemplo, se describe cómo utilizar la consola de Amazon EMR para agregar un paso de Pig a un clúster.

Para enviar un paso de Pig

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com /emr
2. Elija Crear clúster para crear un clúster con Pig instalado. Para obtener información sobre cómo crear un clúster, consulte Planificar y configurar un clúster de Amazon EMR.
3. Para abrir un terminal y SSH en el nodo maestro del clúster, siga los pasos descritos en Conexión al nodo maestro mediante SSH. Una vez hecho esto, ejecute los siguientes pasos.

   ```
   sudo mkdir -p /home/hadoop/lib/pig/
   sudo aws s3 cp s3://elasticmapreduce/libs/pig/0.3/piggybank-0.3-amzn.jar /home/
   hadoop/lib/pig/piggybank.jar
   ```

4. En la consola, haga clic en Lista de clústeres y seleccione el nombre del clúster que ha creado.
5. Desplácese hasta la sección Steps (Pasos) y amplíela; a continuación, elija Add step (Añadir paso).
6. En el cuadro de diálogo Add Step (Añadir paso):

   - En Step type (Tipo de paso), elija Pig program (Programa de Pig).
   - En Name (Nombre), acepte el nombre predeterminado (Programa de Pig) o escriba un nombre nuevo.
   - En Script S3 location (Ubicación en S3 del script), escriba la ubicación del script de Pig. Por ejemplo: **`s3://elasticmapreduce/samples/pig-apache/do-reports2.pig`**.
   - En Input S3 location (Ubicación en S3 de la entrada), escriba la ubicación de los datos de entrada. Por ejemplo: **`s3://elasticmapreduce/samples/pig-apache/input`**.
   - En Ubicación en S3 de la salida, escriba o busque el nombre del bucket de salida de Amazon S3.
   - En Arguments (Argumentos), deje el campo en blanco.
   - En Action on failure (Acción sobre el error), acepte la opción predeterminada, Continue (Continuar).

7. Elija Agregar. El paso aparece en la consola con el estado Pending (Pendiente).
8. El estado del paso cambia de Pending (Pendiente) a Running (En ejecución) y a Completed (Completado) a medida que se ejecuta. Para actualizar el estado, elija el icono Refresh (Actualizar) situado encima de la columna Actions (Acciones). Cuando haya completado el paso, compruebe su bucket de Amazon S3 para confirmar que los archivos de salida de su paso de Pig estén ahí.

Enviar trabajo de Pig mediante la AWS CLI

Para enviar un Pig, haga clic en AWS CLI

Cuando lance un clúster mediante AWS CLI, utilice el `--applications` parámetro para instalar Pig. Para enviar un paso de Pig, utilice el parámetro `--steps`.

1. Para lanzar un clúster con Pig instalado, escriba el siguiente comando *myKey* y sustituya y *amzn-s3-demo-bucket/* por el nombre de su EC2 key pair y bucket de Amazon S3.

```
aws emr create-cluster \
--name "Test cluster" \
--log-uri s3://amzn-s3-demo-bucket/ \
--release-label emr-5.36.2 \
--applications Name=Pig \
--use-default-roles \
--ec2-attributes KeyName=myKey \
--instance-type m5.xlarge \
--instance-count 3
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos secundarios. Todos los nodos utilizan el tipo de instancia especificado en el comando.

> **Note**
>
> Si no ha creado con anterioridad el rol de servicio de EMR predeterminado y el perfil de EC2 instancia, escriba `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

2. Para enviar un paso de Pig, introduzca el siguiente comando *myClusterId* y *amzn-s3-demo-bucket* reemplace por el ID y el nombre del clúster de su bucket de Amazon S3.

```
aws emr add-steps \
--cluster-id myClusterId \
--steps Type=PIG,Name="Pig Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/pig-apache/do-reports2.pig,-p,INPUT=s3://elasticmapreduce/
samples/pig-apache/input,-p,OUTPUT=s3://amzn-s3-demo-bucket/pig-apache/output]
```

   Este comando devolverá un ID de paso, que puede usar para comprobar el `State` de su paso.

3. Consulte el estado del paso con el comando `describe-step`.

```
aws emr describe-step --cluster-id myClusterId --step-id s-1XXXXXXXXXXA
```

   A medida que el paso se ejecuta, el `State` cambia de `PENDING` a `RUNNING` a `COMPLETED`. Cuando haya completado el paso, compruebe su bucket de Amazon S3 para confirmar que los archivos de salida de su paso de Pig estén ahí.

Para más información sobre el uso de comandos de Amazon EMR en la AWS CLI, consulte la Referencia de AWS CLI comandos.

Llamada a funciones definidas por el usuario desde Pig

Pig ofrece la posibilidad de llamar a funciones definidas por el usuario (UDFs) desde los scripts de Pig. Puede hacerlo para implementar el procesamiento personalizado que utilizar en sus scripts de Pig. Los lenguajes compatibles actualmente son Java, Python/Jython y JavaScript (aunque el JavaScript soporte aún es experimental).

En las secciones siguientes se describe cómo registrar sus funciones con Pig para que pueda llamarlos desde el shell de Pig o desde scripts de Pig. Para más información sobre el uso UDFs con Pig, consulte la documentación de Pig correspondiente a su versión de Pig.

Llamar a archivos de JAR desde Pig

Puede utilizar archivos JAR personalizados con Pig mediante el comando `REGISTER` en su script de Pig. El archivo JAR es un sistema de archivos local o remoto como, por ejemplo, Amazon S3. Cuando el script de Pig se ejecuta, Amazon EMR descarga el archivo JAR automáticamente en el nodo maestro y, a continuación, carga el archivo JAR en la caché distribuida de Hadoop. De esta forma, todas las instancias del clúster utilizan automáticamente el archivo JAR según sea necesario.

Para utilizar archivos JAR con Pig

1. Cargue su archivo JAR personalizado en Amazon S3.
2. Utilice el comando `REGISTER` en su script de Pig para especificar el bucket en Amazon S3 del archivo JAR personalizado.

```
REGISTER s3://amzn-s3-demo-bucket/path/mycustomjar.jar;
```

Llamada a scripts de Python/Jython desde Pig

Puede registrar scripts de Python con Pig y, a continuación, llamar a las funciones en dichos scripts desde la shell de Pig o un script de Pig. Esto se hace especificando la ubicación del script con la palabra clave `register`.

Dado que Pig está escrito en Java, utiliza el motor de scripts Jython para analizar scripts de Python. Para obtener más información acerca de Jython, visite http://www.jython.org/.

Para llamar a un script de Python/Jython desde Pig

1. Escriba un script de Python y cargue el script en una ubicación en Amazon S3. Debería ser un bucket propiedad de la misma cuenta que crea el clúster de Pig o que tenga permisos definidos de modo que la cuenta que creó el clúster pueda acceder al mismo. En este ejemplo, el script se carga en s3://amzn-s3-demo - bucket/pig/python
2. Iniciar un clúster de Pig. Si va a acceder a Pig desde el shell de Grunt, ejecute un clúster interactivo. Si está ejecutando comandos de Pig desde un script, inicie un clúster de Pig con scripts. Este ejemplo inicia un clúster interactivo. Para obtener más información sobre crear un clúster de Pig, consulte Enviar trabajo de Pig.

3. Para un clúster interactivo, utilice SSH para conectarse al nodo principal y ejecutar el Grunt shell. Para obtener más información, consulte cómo Conectar con el nodo maestro mediante SSH.
4. Ejecute el shell de Grunt para Pig escribiendo `pig` en la línea de comando:

```
pig
```

5. Registre la biblioteca de Jython y el script de Python con Pig mediante la palabra clave `register` en el símbolo del sistema de Grunt, tal y como se muestra en el comando siguiente, donde puede especificar la ubicación del script en Amazon S3:

```
grunt> register 'lib/jython.jar';
grunt> register 's3://amzn-s3-demo-bucket/pig/python/myscript.py' using jython as
 myfunctions;
```

6. Cargar los datos de entrada. En el siguiente ejemplo, se carga una entrada desde una ubicación de Amazon S3:

```
grunt> input = load 's3://amzn-s3-demo-bucket/input/data.txt' using TextLoader as
 (line:chararray);
```

7. Ahora puede llamar a las funciones en su script desde dentro de Pig haciendo referencia a ellas utilizando `myfunctions`:

```
grunt> output=foreach input generate myfunctions.myfunction($1);
```

Historial de versiones de Pig

En la siguiente tabla, se muestra la versión de Pig incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Pig

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-7.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-7.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-7.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-5.36.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-7.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.0.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.8.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.36.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.36.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.35.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.3.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.2.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.1.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-6.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.34.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.33.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.33.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.32.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.32.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.31.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.31.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.30.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.30.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.30.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.29.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.28.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.28.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.27.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.27.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.26.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.25.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.24.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.24.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.23.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.23.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.22.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.21.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.20.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.20.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.19.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.19.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.18.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.18.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.17.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.17.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.17.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.16.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.16.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.15.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.14.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.13.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.12.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.12.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.11.4	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.11.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.3	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.8.2	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.8.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.7.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.7.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.6.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.6.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.4	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.5.3	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.2	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.5.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.4.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.4.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.3.2	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.3.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.3.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.3	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.2.2	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.2.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.1.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.1.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.3	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.0.2	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.1	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-5.0.0	0.160	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-4.9.6	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.9.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.9.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.8.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.8.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.7.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.6.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.6.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.5.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.4.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.3.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.2.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Etiqueta de versión de Amazon EMR	Versión de Pig	Componentes instalados con Pig
emr-4.1.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.0.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-distcp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Presto

> **Important**
>
> Presto, la versión anterior de Trino, todavía está disponible para su uso con Amazon EMR. Sin embargo, recomendamos encarecidamente utilizar Trino en el futuro con Amazon EMR. Tenga en cuenta también que Trino y Presto no se pueden ejecutar simultáneamente en el mismo clúster. Para obtener más información, consulte Trino.

Presto es un motor de consultas SQL rápido diseñado para realizar consultas de análisis interactivos en grandes conjuntos de datos de diferentes orígenes. Para obtener más información, consulte el sitio web de Presto. Presto se incluye en la versión 5.0.0 de Amazon EMR y posteriores. Las versiones anteriores incluyen Presto como una aplicación de entorno aislado. Para obtener más información, consulte Versiones de lanzamiento de Amazon EMR 4.x. Las versiones 6.1.0 y posteriores de Amazon EMR admiten Trino además de Presto. Para obtener más información, consulte Instalación de PrestoDB.

En la siguiente tabla, se muestra la versión de Presto incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Presto.

Para ver la versión de los componentes instalados con Presto en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Presto para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-7.9.0	Presto 0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
		client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

En la tabla siguiente, se muestra la versión de Presto incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Presto.

Para ver la versión de los componentes instalados con Presto en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Presto para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.15.0	Presto 0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

En la tabla siguiente, se muestra la versión de Presto incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Presto.

Para ver la versión de los componentes instalados con Presto en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Presto para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.36.2	Presto 0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Temas

- Uso de Presto con el Catálogo de datos de AWS Glue
- Uso de S3 Select Pushdown con Presto para mejorar el rendimiento
- Agregar conectores de base de datos
- Uso de SSL/TLS y configuración de LDAPS con Presto en Amazon EMR
- Activación del modo estricto de Presto
- Gestión de la pérdida de instancias de spot en Presto
- Uso del escalado automático de Presto con la retirada estable
- Consideraciones relacionadas con el uso de Presto en Amazon EMR
- Historial de versiones de Presto

Uso de Presto con el Catálogo de datos de AWS Glue

Mediante la versión 5.10.0 o posterior de Amazon EMR, puede especificar el Catálogo de datos de AWS Glue como metaalmacén predeterminado de Hive para Presto. Recomendamos esta

configuración cuando se necesita un metaalmacén persistente o un metaalmacén compartido por diferentes clústeres, servicios, aplicaciones o Cuentas de AWS.

AWS Glue es un servicio completamente administrado de ETL (extracción, transformación y carga) con el que resulta más rentable y sencillo categorizar los datos, limpiarlos, enriquecerlos y moverlos de manera fiable entre distintos almacenes de datos. El Catálogo de datos de AWS Glue proporciona un repositorio de metadatos que unifica toda una variedad de orígenes y formatos de datos, e integra no solo Amazon EMR, sino también Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena y todas las aplicaciones compatibles con el metaalmacén de Apache Hive. AWS Los rastreadores de Glue pueden deducir automáticamente el esquema de los datos de origen en Amazon S3 y almacenar los metadatos asociados en el Catálogo de datos. Para más información acerca del Catálogo de datos, consulte Cómo rellenar el Catálogo de datos de AWS Glue en la AWS Guía para desarrolladores de.

Se aplican cargos separados para AWS Glue. Se paga una tarifa mensual para almacenar los metadatos en el Catálogo de datos y acceder a ellos; una tarifa por hora facturada por minuto para los trabajos de ETL de AWS Glue y el tiempo de ejecución del rastreador; y una tarifa por hora facturada por minuto por cada punto de conexión de desarrollo aprovisionado. El Catálogo de datos le permite almacenar hasta un millón de objetos sin cargo alguno. Si almacena más de un millón de objetos, se le cobrará 1 USD por cada 100 000 objetos que sobrepasen esa cifra. En el Catálogo de datos, se consideran objetos las tablas, las particiones y las bases de datos. Para obtener más información, consulte Precios de Glue.

> **Important**
>
> Si ha creado tablas con Amazon Athena o Amazon Redshift Spectrum antes del 14 de agosto de 2017, las bases de datos y las tablas se almacenan en un catálogo administrado por Athena, que es independiente del Catálogo de datos de Glue. AWS Para integrar Amazon EMR con estas tablas, debe actualizarse al Catálogo de datos de AWS Glue. Para obtener más información, consulte Actualización al catálogo de datos de AWS Glue en la Guía del usuario de Amazon Athena.

Especificación del Catálogo de datos de AWS Glue como metaalmacén

Puede especificar el Catálogo de datos de AWS Glue como metaalmacén mediante la AWS Management Console AWS CLI, la y la API de Amazon EMR. Cuando utiliza la CLI o la API, utiliza la clasificación de configuración de Presto para especificar el Catálogo de datos. Además, con Amazon

EMR 5.16.0 y versiones posteriores, puede utilizar la clasificación de configuración para especificar un Catálogo de datos en otra. Cuenta de AWS Cuando utiliza la consola, puede especificar el Catálogo de datos mediante Opciones avanzadas u Opciones rápidas.

Console

Para especificar el Catálogo de datos de AWS Glue como metaalmacén de Hive con la nueva consola

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Paquete de aplicaciones, seleccione Presto.
4. En la Configuración del Catálogo de datos de AWS Glue, active la casilla de verificación Utilizar para los metadatos de la tabla de Presto.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

CLI

Para especificar el Catálogo de datos de AWS Glue como metaalmacén predeterminado de Hive mediante la AWS CLI

Para ver ejemplos de cómo especificar las siguientes clasificaciones de configuración al crear un clúster, consulte Configuración de aplicaciones.

Versiones 5.16.0 y posteriores de Amazon EMR

- Establezca la propiedad `hive.metastore` en `glue`, tal y como se muestra en el siguiente ejemplo de JSON.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
```

```
]
```

Para especificar un Catálogo de datos en otra Cuenta de AWS, agregue la `hive.metastore.glue.catalogid` propiedad tal y como se muestra en el siguiente ejemplo de JSON. Sustituya *`acct-id`* por la cuenta de AWS del Catálogo de datos. El uso de un Catálogo de datos de otra Cuenta de AWS no está disponible con la versión 5.15.0 y anteriores de Amazon EMR.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Amazon EMR 5.10.0 a 5.15.0

Establezca la propiedad `hive.metastore.glue.datacatalog.enabled` en `true`, tal y como se muestra en el siguiente ejemplo de JSON:

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore.glue.datacatalog.enabled": "true"
    }
  }
]
```

Amazon EMR 6.1.0 y versiones posteriores con PrestoSQL (Trino)

A partir de la versión 6.1.0 de EMR, PrestoSQL también admite Glue como metaalmacén predeterminado de Hive. Utilice la clasificación de configuración `prestosql-connector-hive` para establecer la propiedad `hive.metastore` en `glue`, tal y como se muestra en el siguiente ejemplo de JSON.

Las versiones 6.4.0 y posteriores de Amazon EMR utilizan el nuevo nombre Trino en lugar de PrestoSQL. Si usa Trino, sustituya *prestosql-connector-hive* en la siguiente clasificación de configuración por `trino-connector-hive`.

```
[
  {
    "Classification": "prestosql-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

Para cambiar los metaalmacenes de un clúster de ejecución prolongada, puede establecer manualmente estos valores según corresponda para su versión conectándose al nodo principal, editando los valores de propiedad del archivo `/etc/presto/conf/catalog/hive.properties` directamente y reiniciando el servidor de Presto (`sudo restart presto-server`). Si utiliza este método con Amazon EMR 5.15.0 y versiones anteriores, asegúrese de que `hive.table-statistics-enabled` esté establecido en `false`. Esta configuración no es necesaria cuando se utilizan las versiones 5.16.0 y posteriores; no obstante, no se admiten las estadísticas de tablas y particiones.

Permisos de IAM

El perfil de EC2 instancia de un clúster debe tener permisos de IAM para las acciones de AWS Glue. Además, si se habilita el cifrado para objetos del Catálogo de datos de AWS Glue, el rol también debe poder cifrar, descifrar y generar la AWS KMS key utilizada para el cifrado.

Permisos para las acciones de AWS Glue

Si utiliza el perfil de EC2 instancia predeterminado para Amazon EMR, no se requiere ninguna acción. La política `AmazonElasticMapReduceforEC2Role` gestionada que se adjunta a la `EMR_EC2_DefaultRole` permite todas las acciones de AWS Glue necesarias. Sin embargo, si especifica permisos y un perfil de EC2 instancia personalizados, debe configurar las acciones de AWS Glue adecuadas. Utilice la política administrada `AmazonElasticMapReduceforEC2Role` como punto de partida. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia) en la Guía de administración de Amazon EMR.

Permisos para cifrar y descifrar el Catálogo de datos de AWS Glue

El perfil de instancia necesita permiso para cifrar y descifrar los datos con su clave. No es necesario configurar estos permisos si se cumplen las dos instrucciones siguientes:

- Para habilitar el cifrado de los objetos del Catálogo de datos de AWS Glue, utilice claves administradas para AWS Glue.
- Utilice un clúster que esté en la Cuenta de AWS misma que el Catálogo de datos de AWS Glue.

De lo contrario, debe agregar la siguiente instrucción a la política de permisos adjunta al perfil de EC2 instancia.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Para más información acerca del cifrado del Catálogo de datos de AWS Glue, consulte Cifrado del Catálogo de datos en la AWS Guía para desarrolladores de.

Permisos basados en recursos

Si utiliza AWS Glue junto con Hive, Spark o Presto en Amazon EMR, AWS Glue admite políticas basadas en recursos para controlar el acceso a los recursos del Catálogo de datos. Estos recursos incluyen bases de datos, tablas, conexiones y funciones definidas por el usuario. Para obtener más información, consulte Políticas de recursos de AWS Glue en la Guía para desarrolladores de AWS Glue.

Cuando utilice políticas basadas en recursos para limitar el acceso a AWS Glue desde Amazon EMR, la entidad principal que especifique en la política de permisos debe ser el ARN del rol asociado al perfil de EC2 instancia que se especifica al crear un clúster. Por ejemplo, para una política basada en recursos adjunta a un catálogo, puede especificar el ARN del rol *EMR_EC2_DefaultRole* como rol de servicio predeterminado para las EC2 instancias de clúster`Principal`, con el formato que se muestra en el siguiente ejemplo:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*Puede ser diferente del ID de la cuenta de AWS Glue. Esto permite el acceso desde clústeres de EMR en diferentes cuentas. Puede especificar varias entidades principales, cada una de ellas desde una cuenta diferente.

Consideraciones a la hora de utilizar el Catálogo de datos de AWS Glue

Tenga en cuenta los siguientes aspectos al utilizar el Catálogo de datos de AWS Glue como metaalmacén con Presto:

- No se admite el cambio de nombre de las tablas desde AWS Glue.
- Al crear una tabla de Hive sin especificar un valor de `LOCATION`, los datos de la tabla se almacenan en la ubicación especificada por la propiedad `hive.metastore.warehouse.dir`. De manera predeterminada, se trata de una ubicación en HDFS. Si otro clúster necesita acceder a la tabla, se produce un error a menos que tenga los permisos adecuados para el clúster que creó la tabla. Además, dado que el almacenamiento en HDFS es transitorio, si el clúster termina, se pierden los datos de la tabla y, por lo tanto, será necesario volver a crearla. Se recomienda que especifique una `LOCATION` en Amazon S3 cuando cree una tabla de Hive mediante AWS Glue. Como alternativa, puede utilizar la clasificación de configuración `hive-site` para especificar una ubicación en Amazon S3 para `hive.metastore.warehouse.dir`, que se aplica a todas las tablas de Hive. Si se crea una tabla en una ubicación de HDFS y el clúster que la creó sigue ejecutándose, puede actualizar la ubicación de la tabla a Amazon S3 desde AWS Glue. Para obtener más información, consulte Cómo trabajar con tablas en la consola de AWS Glue en la Guía para desarrolladores de AWS Glue.
- No se admiten valores de partición que contengan comillas y apóstrofes, por ejemplo, `PARTITION (owner="Doe's").`
- Las estadísticas de columnas son compatibles con emr-5.31.0 y versiones posteriores.

- No se admite el uso de la autorización de Hive. Como alternativa, considere la posibilidad de utilizar políticas basadas en recursos de AWS Glue. Para más información, consulta Uso de políticas basadas en recursos para el acceso de Amazon EMR a Catálogo de datos de Glue AWS.

Uso de S3 Select Pushdown con Presto para mejorar el rendimiento

> **Important**
>
> Amazon S3 Select ya no está disponible para los nuevos clientes. Los clientes actuales de Amazon S3 Select pueden seguir utilizando la característica de la forma habitual. Más información

Con la versión 5.18.0 y posteriores de Amazon EMR, puede usar S3 Select Pushdown con Presto en Amazon EMR. Esta característica permite a Presto "delegar" el trabajo de computación de las operaciones de proyección (por ejemplo, `SELECT`) y de las operaciones de predicado (por ejemplo, `WHERE`) en Amazon S3. Esto permite que las consultas recuperen únicamente los datos necesarios desde Amazon S3, lo que puede mejorar el rendimiento y reducir la cantidad de datos transferidos entre Amazon EMR y Amazon S3 en algunas aplicaciones.

¿Es S3 Select Pushdown adecuado para mi aplicación?

Le recomendamos que realice un análisis comparativo sus aplicaciones con y sin S3 Select Pushdown para ver si su uso puede ser adecuado para su aplicación.

Utilice las siguientes directrices para determinar si la aplicación es candidata a utilizar S3 Select:

- La consulta filtra más de la mitad del conjunto de datos original.
- Los predicados del filtro de consulta utilizan columnas que tienen un tipo de datos compatible con Presto y con S3 Select. Los tipos de datos de marca de tiempo, real y doble no son compatibles con S3 Select Pushdown. Le recomendamos que utilice el tipo de datos decimal para los datos numéricos. Para obtener más información acerca de los tipos de datos compatibles con S3 Select, consulte Tipos de datos en la Guía del usuario de Amazon Simple Storage Service.
- La conexión de red entre Amazon S3 y el clúster de Amazon EMR tiene una buena velocidad de transferencia y ancho de banda disponible. Amazon S3 no comprime las respuestas HTTP, por

lo que es probable que el tamaño de la respuesta aumente en el caso de los archivos de entrada comprimidos.

Consideraciones y limitaciones

- Solo se admiten los objetos almacenados en formato CSV. Los objetos se pueden estar sin comprimir o comprimidos con gzip o bzip2.
- No se admite la propiedad `AllowQuotedRecordDelimiters`. Si se especifica esta propiedad, se producirá un error en la consulta.
- No se admiten el cifrado del servidor de Amazon S3 con claves de cifrado proporcionadas por el cliente (SSE-C) ni el cifrado del cliente.
- S3 Select Pushdown no sustituye al uso de formatos de archivo comprimido o en columnas como ORC o Parquet.
- Amazon S3 Select no admite consultas entre regiones para Presto y Trino.

Habilitar S3 Select Pushdown con PrestoDB o Trino

Para habilitar S3 Select Pushdown para PrestoDB en Amazon EMR, utilice la clasificación de configuración `presto-connector-hive` para establecer `hive.s3select-pushdown.enabled` en `true`, tal y como se muestra en el siguiente ejemplo. Para obtener más información, consulte Configuración de aplicaciones. El valor de hive.s3select-pushdown.max-connections también se debe establecer. Para la mayoría de las aplicaciones, la configuración predeterminada, *500*, debería ser suficiente. Para obtener más información, consulte Descripción y ajuste de hive.s3select-pushdown.max-connections más abajo.

Para PrestoSQL en las versiones 6.1.0 a 6.3.0 de EMR, sustituya *presto-connector-hive* en el siguiente ejemplo por `prestosql-connector-hive`.

Las versiones 6.4.0 y posteriores de Amazon EMR utilizan el nuevo nombre Trino en lugar de PrestoSQL. Si usa Trino, sustituya *presto-connector-hive* en el ejemplo siguiente por `trino-connector-hive`

```
[
    {
        "classification": "presto-connector-hive",
        "properties": {
            "hive.s3select-pushdown.enabled": "true",
```

```
            "hive.s3select-pushdown.max-connections": "500"
        }
    }
]
```

Descripción y ajuste de hive.s3select-pushdown.max-connections

De forma predeterminada, Presto utiliza EMRFS como sistema de archivos. El ajuste `fs.s3.maxConnections` de la clasificación de configuración `emrfs-site` especifica el número máximo permitido de conexiones de cliente a Amazon S3 a través de EMRFS para Presto. El valor predeterminado es 500. S3 Select Pushdown no utiliza EMRFS al obtener acceso a Amazon S3 para las operaciones de predicado. En este caso, el valor de `hive.s3select-pushdown.max-connections` determina el número máximo de conexiones de cliente permitidas para dichas operaciones desde los nodos de trabajo. Sin embargo, para todas las solicitudes a Amazon S3 iniciadas por Presto y que no se delegan, como, por ejemplo, las operaciones GET, se sigue teniendo en cuenta el valor de `fs.s3.maxConnections`.

Si la aplicación muestra el error "Timeout waiting for connection from pool" (Tiempo de espera esperando la conexión desde el grupo), incremente los valores de `hive.s3select-pushdown.max-connections` y `fs.s3.maxConnections`.

Agregar conectores de base de datos

Puede utilizar clasificaciones de configuración para configurar las propiedades del conector JDBC al crear un clúster. Las clasificaciones de configuración comienzan por `presto-connector`, por ejemplo, `presto-connector-postgresql`. Las clasificaciones de configuración disponibles varían en función de la versión de Amazon EMR. Para ver las clasificaciones de configuración disponibles con la versión más reciente, consulte the section called “Clasificaciones de configuración de la versión 5.36.2” para Amazon EMR 5.36.2. Si utiliza una versión diferente de Amazon EMR, consulte Versiones de lanzamiento de Amazon EMR 5.x para conocer las clasificaciones de configuración. Para más información acerca de las propiedades que se pueden configurar con cada conector, consulte https://prestodb.io/docs/current/connector.html.

Example : configuración de un clúster con el conector JDBC de PostgreSQL

Para lanzar un clúster con el conector de PostgreSQL instalado y configurado, primero debe crear un archivo JSON que especifique la clasificación de configuración (por ejemplo, `myConfig.json`) con el siguiente contenido, y guarde localmente.

Sustituya las propiedades de conexión según sea apropiado para su configuración, tal y como se muestra en el tema Conector de PostgreSQL de la documentación de Presto.

```
[
  {
    "Classification": "presto-connector-postgresql",
    "Properties": {
      "connection-url": "jdbc:postgresql://example.net:5432/database",
      "connection-user": "MYUSER",
      "connection-password": "MYPASS"
    },
    "Configurations": []
  }
]
```

Al crear el clúster, haga referencia a la ruta del archivo JSON mediante la opción `--configurations`, tal y como se muestra en el siguiente ejemplo, donde `myConfig.json` se encuentra en el mismo directorio en el que se ejecuta el comando:

```
aws emr create-cluster --name PrestoConnector --release-label emr-5.36.2 --instance-
type m5.xlarge \
--instance-count 2 --applications Name=Hadoop Name=Hive Name=Pig Name=Presto \
--use-default-roles --ec2-attributes KeyName=myKey \
--log-uri s3://amzn-s3-demo-bucket/logs --enable-debugging \
--configurations file://myConfig.json
```

Uso de SSL/TLS y configuración de LDAPS con Presto en Amazon EMR

Con Amazon EMR versión 5.6.0 y posteriores, puede habilitar SSL/TLS para ayudar a proteger las comunicaciones internas entre los nodos de Presto. Para ello, se establece una configuración de seguridad para el cifrado en tránsito. Para obtener más información, consulte Opciones de cifrado y Uso de configuraciones de seguridad para definir la seguridad del clúster en la Guía de administración de Amazon EMR.

Cuando se utiliza una configuración de seguridad con el cifrado en tránsito, Amazon EMR hace lo siguiente para Presto:

- Distribuye los artefactos de cifrado o los certificados que especifique para el cifrado en tránsito a lo largo del clúster de Presto. Para obtener más información, consulte Proporcionar certificados para cifrado de datos en tránsito.
- Establece las siguientes propiedades utilizando la clasificación de configuración `presto-config`, que se corresponde con el archivo `config.properties` de Presto:
  - Establece `http-server.http.enabled` en `false` en todos los nodos, lo que deshabilita HTTP en favor de HTTPS. Esto requiere que proporcione certificados que funcionen para DNS públicos y privados al configurar la configuración de seguridad para el cifrado en tránsito. Una forma de hacerlo es utilizar certificados SAN (nombre alternativo del sujeto) que admitan varios dominios.
  - Establece los valores de `http-server.https.*`. Para conocer los detalles de configuración, consulte LDAP Authentication en la documentación de Presto.

Además, con Amazon EMR versión 5.10.0 y posteriores, puede configurar la autenticación LDAP para las conexiones de cliente con el coordinador de Presto mediante HTTPS. Esta configuración utiliza LDAP seguro (LDAPS). TLS debe estar habilitado en su servidor LDAP y el clúster de Presto debe utilizar una configuración de seguridad con el cifrado de datos en tránsito habilitado. Se requiere configuración adicional. Las opciones de configuración son diferentes en función de la versión de Amazon EMR que utilice. Para obtener más información, consulte Uso de la autenticación LDAP para Presto en Amazon EMR.

Presto en Amazon EMR utiliza el puerto 8446 para HTTPS interno de forma predeterminada. El puerto utilizado para la comunicación interna debe ser el mismo puerto utilizado para el acceso HTTPS del cliente al coordinador de Presto. La propiedad `http-server.https.port` de la configuración de clasificación `presto-config` especifica el puerto.

Uso de la autenticación LDAP para Presto en Amazon EMR

Siga los pasos de esta sección para configurar LDAP. Examine cada paso para ver ejemplos y enlaces a información adicional.

Pasos para configurar la autenticación de LDAP

- Paso 1: recopilar información sobre su servidor LDAP y copiar el certificado de servidor en Amazon S3
- Paso 2: Definir una configuración de seguridad
- Paso 3: Crear una configuración JSON con propiedades de Presto para LDAP

- Paso 4: crear el script para copiar el certificado de servidor LDAP y cargarlo en Amazon S3
- Paso 5: Crear el clúster

Paso 1: recopilar información sobre su servidor LDAP y copiar el certificado de servidor en Amazon S3

Necesitará la información y los elementos de la siguiente sección de su servidor LDAP para configurar la autenticación de LDAP.

La dirección IP o el nombre de host del servidor LDAP

El coordinador de Presto en el nodo maestro de Amazon EMR debe poder conectarse al servidor LDAP en la dirección IP o el nombre de host especificados. De forma predeterminada, Presto se comunica con el servidor LDAP mediante LDAPS a través del puerto 636. Si su implementación de LDAP requiere un puerto personalizado, puede especificarlo mediante la propiedad `ldap.url` con Amazon EMR 5.16.0 o versiones posteriores, o utilizar `authentication.ldap.url` con versiones anteriores. Sustituya el puerto personalizado de 636 que se muestra en los ejemplos de clasificación de configuración `presto-config` en Paso 3: Crear una configuración JSON con propiedades de Presto para LDAP. Asegúrese de que todos los firewalls y grupos de seguridad permiten el tráfico entrante y saliente en el puerto 636 (o en el puerto personalizado) y también en el puerto 8446 (o en el puerto personalizado), que se utiliza para las comunicaciones internas del clúster.

El certificado del servidor LDAP

Debe cargar el archivo de certificado en una ubicación segura en Amazon S3. Para obtener más información, consulte ¿Cómo puedo cargar archivos y carpetas en un bucket de S3? en la Guía del usuario de Amazon Simple Storage Service. Debe crear una acción de arranque que copie este certificado de Amazon S3 en cada nodo del clúster cuando se lance el clúster. En Paso 4: crear el script para copiar el certificado de servidor LDAP y cargarlo en Amazon S3. El certificado de ejemplo es *`s3://amzn-s3-demo-bucket/ldap_server.crt`*.

La configuración del servidor LDAP para la conexión anónima

Si la conexión anónima está deshabilitada en PrestoDB, necesita el ID de usuario (UID) y la contraseña de una cuenta con permisos para conectarse al servidor LDAP, de modo que el servidor de PrestoDB pueda establecer una conexión. El UID y la contraseña se especifican con las propiedades `internal-communication.authentication.ldap.user` e `internal-communication.authentication.ldap.password` en la clasificación de configuración

`presto-config`. Amazon EMR 5.10.0 no admite esta configuración, por lo que la conexión anónima debe permitirse en el servidor LDAP cuando se utiliza esta versión.

Tenga en cuenta que Trino no requiere la configuración de enlace anónimo.

Para obtener el estado de la conexión anónima en el servidor LDAP

- Utilice el comando ldapwhoami desde un cliente Linux, tal y como se muestra en el ejemplo siguiente:

  ```
  ldapwhoami -x -H ldaps://LDAPServerHostNameOrIPAddress
  ```

  Si la conexión anónima no está permitida, el comando devuelve lo siguiente:

  ```
  ldap_bind: Inappropriate authentication (48)
  additional info: anonymous bind disallowed
  ```

Para verificar que una cuenta tiene permisos para un servidor LDAP que utiliza la autenticación sencilla

- Utilice el comando ldapwhoami desde un cliente Linux, tal y como se muestra en el ejemplo siguiente. El ejemplo utiliza un usuario ficticio almacenado en un servidor Open LDAP que se ejecuta en una EC2 instancia con el nombre de host ficticio. *presto ip-xxx-xxx-xxx-xxx.ec2.internal* El usuario está asociado a la unidad organizativa (OU) *admins* y a la contraseña: *123456*

  ```
  ldapwhoami -x -w "123456" -D uid=presto,ou=admins,dc=ec2,dc=internal -H ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal
  ```

  Si la cuenta es válida y tiene los permisos adecuados, el comando devuelve:

  ```
  dn:uid=presto,ou=admins,dc=ec2,dc=internal
  ```

Las configuraciones de ejemplo de Paso 3: Crear una configuración JSON con propiedades de Presto para LDAP incluyen esta cuenta por razones de claridad, con la excepción del ejemplo de 5.10.0, en el que no se admite. Si el servidor LDAP utiliza la conexión anónima, quite el par

de nombre-valor `internal-communication.authentication.ldap.user` e `internal-communication.authentication.ldap.password`.

El nombre distintivo (DN) de LDAP para los usuarios de Presto

Al especificar la configuración de LDAP para Presto, se especifica un patrón de `${USER}` enlace que consiste en una unidad organizativa (OU) y componentes de dominio adicionales (). DCs Presto sustituye `${USER}` por el ID de usuario (UID) real de cada usuario durante la autenticación de contraseñas para que coincida con el nombre distintivo (DN) que este patrón de conexión especifica. Necesita la propiedad a la OUs que pertenecen los usuarios aptos y sus DCs Por ejemplo, para permitir que los usuarios de la unidad organizativa `admins` del dominio `corp.example.com` se autentiquen en Presto, debe especificar `${USER},ou=admins,dc=corp,dc=example,dc=com` como el patrón de conexión del usuario.

> **Note**
>
> Cuando utilice AWS CloudFormation, tendrá que utilizar la función Fn: :Sub para `${USER}` sustituirla por el ID de usuario (UID) real. Para obtener más información, consulte el tema Fn::Sub en la Guía del usuario de AWS CloudFormation .

Cuando utilice Amazon EMR 5.10.0, solo puede especificar un patrón de este tipo. Con Amazon EMR 5.11.0 o versiones posteriores, puede especificar varios patrones separados por un signo de dos puntos (:). Los usuarios que intentan autenticarse en Presto se cotejan con el primer patrón, después con el segundo y así sucesivamente. Para ver un ejemplo, consulta Paso 3: Crear una configuración JSON con propiedades de Presto para LDAP.

Paso 2: Definir una configuración de seguridad

Cree una configuración de seguridad con el cifrado en tránsito habilitado. Para obtener más información, consulte Creación de una configuración de seguridad en la Guía de administración de Amazon EMR. Los artefactos de cifrado que proporcionó al configurar el cifrado en tránsito se utilizan para cifrar la comunicación interna entre los nodos de Presto. Para obtener más información, consulte Proporcionar certificados para cifrado de datos en tránsito. El certificado de servidor LDAP se utiliza para autenticar las conexiones cliente con el servidor de Presto.

Paso 3: Crear una configuración JSON con propiedades de Presto para LDAP

Se utiliza la clasificación de configuración `presto-config` para establecer las propiedades de Presto para LDAP. El formato y el contenido de `presto-config` varían ligeramente según la versión de lanzamiento de Amazon EMR y la instalación de Presto (PrestoDB o Trino). Más adelante en esta sección, se proporcionan ejemplos de las diferencias de configuración. Para obtener más información, consulte Configuración de aplicaciones.

En los siguientes pasos, se presupone que se guardan los datos de JSON en un archivo,*MyPrestoConfig.json*. Si utiliza la consola, cargue el archivo en un lugar seguro en Amazon S3 para que pueda hacer referencia a él cuando cree el clúster. Si usa el AWS CLI, puede hacer referencia al archivo localmente.

Example Amazon EMR 6.1.0 y versiones posteriores con PrestoSQL (Trino)

En el siguiente ejemplo, se utiliza el nombre de host de LDAP de Paso 1: recopilar información sobre su servidor LDAP y copiar el certificado de servidor en Amazon S3 para la autenticación en el servidor LDAP para la conexión. Se especifican dos patrones de conexión de usuario, que indican que los usuarios dentro de las unidades organizativas `admins` y `datascientists` del servidor LDAP reúnen los requisitos para la autenticación en el servidor de Trino como usuarios. Los patrones de conexión están separados por un signo de dos puntos (`:`).

Las versiones 6.4.0 y posteriores de Amazon EMR utilizan el nuevo nombre Trino en lugar de PrestoSQL. Si usa Trino, sustituya *`prestosql-config`* en la siguiente clasificación de configuración por `trino-config` y *`prestosql-password-authenticator`* por `trino-password-authenticator`.

```
[
   {
      "Classification":"prestosql-config",
      "Properties":{
         "http-server.authentication.type":"PASSWORD"
      }
   },
   {
      "Classification":"prestosql-password-authenticator",
      "Properties":{
         "password-authenticator.name":"ldap",
         "ldap.url":"ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
         "ldap.user-bind-pattern": "uid=${USER},ou=admins,dc=ec2,dc=internal:uid=
${USER},ou=datascientists,dc=ec2,dc=internal"
```

```
      }
   }
]
```

Example Versiones 5.16.0 y posteriores de Amazon EMR

En el siguiente ejemplo se utiliza el ID de usuario y la contraseña de LDAP y el nombre de host de LDAP Paso 1: recopilar información sobre su servidor LDAP y copiar el certificado de servidor en Amazon S3 para la autenticación en el servidor LDAP para la conexión. Se especifican dos patrones de conexión de usuario, que indican que los usuarios dentro de las unidades organizativas `admins` y `datascientists` del servidor LDAP reúnen los requisitos para la autenticación en el servidor de Presto como usuarios. Los patrones de conexión están separados por un signo de dos puntos (`:`).

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "PASSWORD"
                }
        },
        {
                "Classification": "presto-password-authenticator",
                "Properties": {
                        "password-authenticator.name": "ldap",
                        "ldap.url": "ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
                        "ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
                }
        }]
```

Example Amazon EMR 5.11.0 a 5.15.0

El formato de la clasificación de configuración `presto-config` es ligeramente diferente para estas versiones. En el siguiente ejemplo se especifican los mismos parámetros que en el ejemplo anterior.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
```

```
                        "authentication.ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
                }
        }]
```

Example Amazon EMR 5.10.0

Amazon EMR 5.10.0 solo admite la conexión anónima, por lo que se omiten estas entradas. Además, solo se puede especificar un patrón de conexión.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
                        "ldap.user-bind-pattern": "uid=
${USER},ou=prestousers,dc=ec2,dc=internal"
                }
        }]
```

Paso 4: crear el script para copiar el certificado de servidor LDAP y cargarlo en Amazon S3

Cree un script que copie el archivo de certificado en cada nodo del clúster y lo añada al almacén de claves. Cree el script con un editor de texto, guárdelo y, a continuación, cárguelo en Amazon S3. EnPaso 5: Crear el clúster, se hace referencia al archivo de script como*s3://amzn-s3-demo-bucket/LoadLDAPCert.sh*.

El siguiente script de ejemplo utiliza la contraseña predeterminada del almacén de claves,*changeit*. Le recomendamos que se conecte al nodo principal después de crear el clúster y que cambie la contraseña del almacén de claves con el comando keytool.

```
#!/bin/bash
aws s3 cp s3://amzn-s3-demo-bucket/ldap_server.crt .
sudo keytool -import -keystore /usr/lib/jvm/jre-1.8.0-openjdk.x86_64/lib/security/
cacerts -trustcacerts -alias ldap_server -file ./ldap_server.crt -storepass changeit -
noprompt
```

Paso 5: Crear el clúster

Al crear el clúster, debe especificar Presto y demás aplicaciones que desee que Amazon EMR instale. Los siguientes ejemplos también hacen referencia a las propiedades de clasificación de configuración de un archivo JSON, pero también puede especificar la clasificación de configuración directamente.

Para crear un clúster de Presto con autenticación LDAP mediante la consola de Amazon EMR

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. Elija Presto junto con las demás aplicaciones que desee que Amazon EMR instale y, en Configuración de software, seleccione un valor en Versión para la versión de Amazon EMR que desee utilizar. La autenticación LDAP solo se admite con Amazon EMR 5.10.0 y versiones posteriores.
4. En Editar configuración de software, elija Cargar JSON desde S3, escriba la ubicación en Amazon S3 del archivo de configuración JSON que creó en Paso 3: Crear una configuración JSON con propiedades de Presto para LDAP y, a continuación, elija Siguiente.
5. Configure el hardware y la conexión en red del clúster y, a continuación, elija Next (Siguiente).
6. Elija Bootstrap Actions (Acciones de arranque). En Add bootstrap action (Añadir acción de arranque), seleccione Custom action (Acción personalizada) y después elija Configure and add (Configurar y añadir).
7. Introduzca un nombre para la acción de arranque, introduzca la ubicación del script en la que la creó (por ejemploPaso 4: crear el script para copiar el certificado de servidor LDAP y cargarlo en Amazon S3, s3://amzn-s3-demo-bucket/Load LDAPCert .sh) y, a continuación, seleccione Añadir.
8. En General Options (Opciones generales), Tags (Etiquetas) y Additional Options (Opciones adicionales), elija la configuración adecuada para su aplicación y después elija Next (Siguiente).
9. Elija Authentication and encryption (Autenticación y cifrado) y, a continuación, seleccione el valor de Security configuration (Configuración de seguridad) correspondiente a la configuración de seguridad que creó en el Paso 2: Definir una configuración de seguridad.
10. Elija otras opciones de seguridad que sean necesarias para su aplicación y, a continuación, elija Create cluster (Crear clúster).

Para crear un clúster de Presto con autenticación LDAP mediante la AWS CLI

- Utilice el comando `aws emr create-cluster`. Como mínimo, especifique la aplicación de Presto y también la clasificación de configuración de Presto, el script de arranque y la configuración de seguridad que creó en los pasos anteriores. El siguiente ejemplo hace referencia al archivo de configuración como un archivo JSON guardado en el mismo directorio en el que se ejecuta el comando. El script de arranque, por otro lado, debe guardarse en Amazon S3. El siguiente ejemplo utiliza `s3://amzn-s3-demo-bucket/LoadLDAPCert.sh`.

  > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster  --applications Name=presto --release-label emr-5.16.0 \
--use-default-roles --ec2-attributes KeyName=MyKeyPair,SubnetId=subnet-1234ab5
 \ --instance-count 3 --instance-type m5.xlarge --region us-west-2 --name
 "MyPrestoWithLDAPAuth" \
--bootstrap-actions Name="Distribute LDAP server cert",Path="s3://amzn-s3-demo-
bucket/LoadLDAPCert.sh" \
--security-configuration MyPrestoLDAPSecCfg --configurations file://
MyPrestoConfig.json
```

Activación del modo estricto de Presto

En determinadas situaciones, las consultas de larga duración pueden generar costos elevados y hacer que Amazon EMR utilice más recursos del clúster. Esto reduce los recursos de otras cargas de trabajo del clúster. Con las versiones 6.8 y posteriores de Amazon EMR, puede utilizar una característica de modo estricto que rechaza o advierte sobre los siguientes tipos de consultas de larga duración:

- Consultas sin predicados en las columnas particionadas que dan como resultado el escaneo de tablas de grandes cantidades de datos
- Consultas con uniones cruzadas entre dos tablas grandes
- Consultas que ordenan un gran número de filas sin límite

Una vez que Presto optimiza completamente el plan de consultas, se ejecuta el modo estricto. Para usar y personalizar el modo estricto según sus necesidades de consulta, puede configurar Presto de las siguientes maneras.

Configuraciones de Presto para el modo estricto

Opción	Descripción	Predeterminado
`strict-mode-enabled`	Activa y desactiva el modo estricto. Un valor de `true` indica que el modo estricto está activado.	`false`
`strict-mode-fail-query`	Rechaza las consultas si el modo estricto detecta consultas probables de larga duración. Si `false`, Amazon EMR solo genera una advertencia.	`false`
`strict-mode-restrictions`	Especifica las restricciones que se aplicarán cuando se active el modo estricto. El modo estricto admite las siguientes restricciones: MANDATORY_PARTITION_PREDICATE, DISALLOW_CROSS_JOIN, and LIMITED_SORT.	MANDATORY_PARTITION_PREDICATE,DISALLOW_CROSS_JOIN, LIMITED_SORT

Para experimentar con el modo estricto, puede anular estas configuraciones o establecerlas como propiedades de sesión cuando utilice el cliente Presto.

Para establecer la configuración en el momento de la creación del clúster con la AWS Management Console

1. Elija Crear clúster y seleccione Amazon EMR versión 6.8.0 y Presto o Trino. Para obtener más información, consulte Instalación de PrestoDB y Trino.

2. Especifique directamente las propiedades de configuración para el modo estricto o cargue un archivo JSON en Amazon S3. Si lo desea, seleccione el catálogo de datos de AWS Glue para su metaalmacén. Especifique la VPC, las subredes, las acciones de arranque, el par de claves y el grupo de seguridad. Elija Crear cluster para crear el clúster.
3. Inicie sesión en el nodo principal del clúster y ejecute `presto-cli` o `trino-cli`.
4. Envíe sus consultas. El modo estricto valida cada consulta y determina si es de larga duración. Según la configuración de `strict-mode-fail-query`, Amazon EMR rechaza la consulta o genera una advertencia.
5. Cuando haya terminado con las consultas, finalice el clúster y elimine los recursos.

Para establecer la configuración en un clúster en funcionamiento con la AWS CLI

1. Inicie sesión en el nodo principal del clúster con AWS CLI y ejecute `presto-cli` o`trino-cli`.
2. Ejecute los siguientes comandos con los valores que desee.

```
set session strict_mode_enabled = true;
set session strict_mode_fail_query = false;
set session strict_mode_restrictions = 'DISALLOW_CROSS_JOIN,LIMITED_SORT';
```

Consideraciones

Cuando utilice el modo estricto, tenga en cuenta lo siguiente:

- En algunos casos, el modo estricto puede rechazar consultas de corta duración que no consuman muchos recursos. Por ejemplo, las consultas en tablas pequeñas no aplican filtros dinámicos ni sustituyen las uniones internas por uniones cruzadas. Esto puede llevar a la consulta a utilizar el predicado de partición obligatorio o a no permitir la unión cruzada. Cuando esto sucede, el modo estricto rechaza la consulta.
- La comprobación del modo estricto solo se aplica a los tipos de consulta SELECT, INSERT, CREATE TABLE AS SELECT y EXPLAIN ANALYZE.
- Solo puede utilizar la restricción en el predicado de partición obligatorio con el conector Hive.

Gestión de la pérdida de instancias de spot en Presto

Con las instancias de spot de Amazon EMR, puede ejecutar cargas de trabajo de macrodatos en la EC2 capacidad sobrante de Amazon a un costo reducido. A cambio de un menor costo, Amazon EC2 puede interrumpir las instancias de spot con una notificación de dos minutos. Al cerrar un nodo, Presto puede tardar hasta 10 minutos en generar un error. Esto provoca demoras innecesarias en los informes de errores y en los posibles reintentos. La terminación rápida es una característica que le permite controlar la forma en que Presto gestiona los nodos terminados.

El trabajo del coordinador de Presto consiste en hacer un seguimiento de todos los nodos de trabajo mediante sondeos periódicos sobre sus estados. Sin una terminación rápida, el coordinador no consulta el estado de cada nodo en el YARN NodeManager . Esto puede provocar un ciclo de reintento prolongado antes de que la consulta falle. Con una terminación rápida, el coordinador de Presto consulta el estado del nodo tan pronto NodeManager como la encuesta no llegue al host. Si se NodeManager muestra que el nodo está inactivo, Presto abandona los reintentos, no se completa la consulta y devuelve un error. `NODE_DECOMMISSIONED`

El siguiente conjunto de parámetros de configuración le permite controlar y personalizar el comportamiento de Presto en caso de que se cierre el nodo.

Configuraciones de Presto para la gestión de errores de los nodos

Opción	Descripción	Predeterminado
`query.remote-task.max-backoff-duration`	El tiempo que el coordinador continúa intentando recuperar el estado de la tarea remota desde los nodos de trabajo.	10 minutos
`query.remote-task.quick-terminate-node-failure`	Activa un error rápido en el nodo si el coordinador no puede llegar al nodo o no puede conectarse con el trabajador que se ejecuta en ese nodo. El valor de `query.remote-task.terminate-on-connect-exception` determina si el coordinador debe llegar	`true`

Opción	Descripción	Predeterminado
	al nodo o conectarse con el trabajador. El nodo no supera la consulta y Amazon EMR lo elimina de la lista de trabajadores disponibles. Cuando esto sucede, no puede usar el nodo para programar consultas nuevas. Si establece este valor en `false`, Presto vuelve a su comportamiento anterior, en el que el coordinador de Presto vuelve a intentar llegar al nodo (para `query.remote-task.max-backoff-duration`) antes de marcar el nodo como no disponible y no llevar a cabo la consulta en curso en el nodo.	
`query.remote-task.terminate-on-connect-exception`	Especifica si Amazon EMR debe ser un nodo si se puede acceder al host, pero el coordinador no puede conectarse al proceso de trabajo del host. Al establecer este valor en `true`, se activa el error de consulta rápida si no se puede acceder al host.	`false`

Uso del escalado automático de Presto con la retirada estable

Las versiones 5.30.0 y posteriores de Amazon EMR incluyen una característica que puede utilizar para establecer un periodo de gracia para determinadas acciones de escalado. El periodo de gracia permite que las tareas de Presto se ejecuten antes de que se termine el nodo debido a una acción de cambio para reducir horizontalmente o a una solicitud de política de escalado automático. Para obtener más información sobre las reglas de escalado, consulte Descripción de las reglas de escalado automático en la Guía de administración de Amazon EMR. El escalado automático con retirada estable de Presto evita que se programen nuevas tareas en un nodo que se está retirando, mientras que, al mismo tiempo, permite que las tareas que ya se están ejecutando se completen antes de que se alcance el tiempo de espera de cierre. Las consultas que se están ejecutando terminarán de ejecutarse antes de que se retire el nodo. El escalado automático no se admite en las flotas de instancias.

Puede controlar el tiempo que tardarán en completarse las tareas de Presto después de recibir una solicitud de cierre con escala automática. De forma predeterminada, el tiempo de espera de cierre de Amazon EMR es de `0` minutos, lo que significa que Amazon EMR termina inmediatamente el nodo y cualquier tarea de Presto que se ejecute en él, si así lo requiere una solicitud de escalado. Para establecer un tiempo de espera más largo para las tareas de Presto en Amazon EMR para permitir que las consultas en ejecución se completen antes de reducir la capacidad de un clúster, utilice la clasificación de configuración `presto-config` para establecer el parámetro `graceful-shutdown-timeout` en un valor de segundos o minutos mayores que cero. Para obtener más información, consulte Configuración de aplicaciones.

Por ejemplo, cuando se aumenta el valor de `graceful-shutdown-timeout` a `"30m"`, se especifica un periodo de 30 minutos. Una vez finalizado el periodo de tiempo de espera de cierre, se aplica un cierre forzoso al nodo marcado para retirada si está esperando a que se completen las tareas de consulta y se produce un error en la consulta. Si las tareas de consulta finalizan en cinco minutos, el nodo marcado para retirada terminará, siempre que otras aplicaciones de YARN hayan completado la ejecución.

Example Ejemplo de configuración de escalado automático con retirada estable de nodos

Reemplace el valor de `graceful-shutdown-timeout` por el número de minutos apropiado para su configuración. No hay ningún valor máximo. El siguiente ejemplo establece un valor de tiempo de espera de `1800` segundos (30 minutos).

```
[
    {
```

```
        "classification": "presto-config",
        "properties": {
            "graceful-shutdown-timeout": "1800s"
        }
    }
]
```

Limitaciones

PrestoDB Graceful Decommission no funciona en clústeres de EMR en los que la conectividad HTTP está deshabilitada, por ejemplo, cuando `http-server.http.enabled` está configurada en `false`. Trino no es compatible en absoluto con Graceful Decommission, independientemente de la configuración `http-server.http.enabled`.

Consideraciones relacionadas con el uso de Presto en Amazon EMR

Tenga en cuenta las siguientes limitaciones al ejecutar Presto en Amazon EMR.

Ejecutable de línea de comandos de Presto

En Amazon EMR, PrestoDB y Trino utilizan el mismo ejecutable de línea de comandos, `presto-cli`, como en el siguiente ejemplo.

```
presto-cli --catalog hive
```

Propiedades de la implementación de Presto no configurables

La versión de Amazon EMR que utilice determina las configuraciones de implementación de Presto disponibles. Para obtener más información sobre estas propiedades de configuración, consulte Implementación de Presto en la documentación de Presto. En la siguiente tabla, se muestran las distintas opciones de configuración de los archivos `properties` de Presto.

Archivos	Configurable
`log.properties`	PrestoDB: versiones 4.0.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `presto-log` .

Archivos	Configurable
`config.properties`	PrestoDB: versiones 4.0.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `presto-config` .
`hive.properties`	PrestoDB: versiones 4.1.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `presto-connector-hive` .
`node.properties`	PrestoDB: versiones 5.6.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `presto-node` .
`jvm.config`	No se puede configurar.

Instalación de PrestoDB

El nombre de la aplicación Presto se sigue utilizando para instalar PrestoDB en los clústeres.

Puede instalar PrestoDB o Trino, pero no puede instalar ambos en un solo clúster. Si se especifican PrestoDB y Trino al intentar crear un clúster, se produce un error tanto en la validación como en la solicitud de creación del clúster.

Configuración de EMRFS y PrestoS3PrestoS3 FileSystem

Con las versiones 5.12.0 y posteriores de Amazon EMR, PrestoDB puede utilizar EMRFS. Para obtener más información, consulte Sistema de archivos de EMR (EMRFS) en la Guía de administración de Amazon EMR. En las versiones anteriores de Amazon EMR, PrestoS3 FileSystem es la única opción de configuración.

Se puede utilizar una configuración de seguridad para configurar el cifrado de los datos de EMRFS en Amazon S3. También puede utilizar roles de IAM para las solicitudes de EMRFS a Amazon S3. Para obtener más información, consulte Descripción de las opciones de cifrado y Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3 en la Guía de administración de Amazon EMR.

> **Note**
>
> Si consulta los datos subyacentes en Amazon S3 con la versión 5.12.0 de Amazon EMR, pueden producirse errores en Presto. Esto se debe a que Presto no recoge los valores de clasificación de configuración de `emrfs-site.xml`. Como solución alternativa, cree un subdirectorio de `emrfs` en `usr/lib/presto/plugin/hive-hadoop2/` y cree un enlace simbólico en `usr/lib/presto/plugin/hive-hadoop2/emrfs` al archivo `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` existente. A continuación, reinicie el proceso presto-server (`sudo presto-server stop` seguido de `sudo presto-server start`).

Puede anular el EMRFS predeterminado y utilizar PrestoS3 en su lugar. FileSystem Para ello, utilice la clasificación de configuración `presto-connector-hive` para establecer `hive.s3-file-system-type` en `PRESTO`, tal como se muestra en el siguiente ejemplo. Para obtener más información, consulte Configuración de aplicaciones.

```
[
   {
      "Classification": "presto-connector-hive",
      "Properties": {
         "hive.s3-file-system-type": "PRESTO"
      }
   }
]
```

Si utiliza PrestoS3, utilice la clasificación de configuración para configurar las FileSystem propiedades de PrestoS3. `presto-connector-hive` FileSystem Para obtener más información sobre las propiedades disponibles, consulte Configuración de Amazon S3 en la sección Conector de Hive de la documentación de Presto. Estas opciones no se aplican a EMRFS.

Configuración predeterminada para la suplantación del usuario final

De forma predeterminada, las versiones 5.12.0 y posteriores de Amazon EMR permiten la suplantación del usuario final para obtener acceso a HDFS. Para obtener más información, consulte Suplantación de usuarios finales en la documentación de Presto. Para cambiar esta configuración con la clasificación de configuración de `presto-config`, establezca la propiedad `hive.hdfs.impersonation.enabled` en `false`.

Puerto predeterminado para la interfaz web de Presto

De forma predeterminada, Amazon EMR configura la interfaz web de Presto en el coordinador de Presto para que utilice el puerto 8889 (para PrestoDB y Trino). Para cambiar el puerto, utilice la clasificación de configuración `presto-config` para establecer la propiedad de `http-server.http.port`. Para obtener más información, consulte Propiedades de configuración en la sección Implementación de Presto de la documentación de Presto.

Problema con la ejecución del bucket de Hive en algunas versiones

Presto versión 152.3 tiene un problema con la ejecución de buckets de Hive que provoca un rendimiento de consultas de Presto notablemente más lento en algunas circunstancias. Las versiones 5.0.3, 5.1.0 y 5.2.0 de Amazon EMR incluyen esta versión de Presto. Para mitigar este problema, utilice la clasificación de configuración `presto-connector-hive` para establecer la propiedad `hive.bucket-execution` en `false`, tal y como se muestra en el siguiente ejemplo.

```
[
   {
      "Classification": "presto-connector-hive",
      "Properties": {
         "hive.bucket-execution": "false"
      }
   }
]
```

Historial de versiones de Presto

En la siguiente tabla, se muestra la versión de Presto incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Presto

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-7.9.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.8.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.7.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
		yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.6.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.5.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-7.4.0	0,287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.3.0	0,285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-7.2.0	0,285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.36.2	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-7.1.0	0,284	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.0.0	0,283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.15.0	0,283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.14.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.13.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.12.0	0,281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.11.1	0,279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.11.0	0,279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.10.1	0,278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.10.0	0,278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.9.1	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.9.0	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.8.1	0,273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.8.0	0,273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.7.0	0,272	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.36.1	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.36.0	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.6.0	0,267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.35.0	0,266	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.5.0	0,261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.4.0	0.254.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.3.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.3.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.2.1	0.238.238.	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.2.0	0.238.238.	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.1.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.1.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.0.1	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-6.0.0	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.34.0	0,261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.33.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.33.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.32.1	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.32.0	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.31.1	0.238.238.	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.31.0	0.238.238.	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.30.2	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.30.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.29.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.28.1	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.28.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.27.1	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.27.0	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.26.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.25.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.24.1	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.24.0	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.23.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.23.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.22.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.21.2	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.21.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.21.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.20.1	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.20.0	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.19.1	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.19.0	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.18.1	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.18.0	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.17.2	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.1	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.17.0	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.16.1	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.16.0	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.15.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.15.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.14.2	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.14.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.14.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.13.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.13.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.12.3	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.2	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.12.1	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.0	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.11.4	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.3	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.11.2	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.11.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.10.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.10.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.9.1	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.9.0	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.8.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.8.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.7.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.7.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.6.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.6.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.4	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.5.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.4.1	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.4.0	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.3.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.3.0	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.2.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.2.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.1.1	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.1.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.3	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Etiqueta de versión de Amazon EMR	Versión de Presto	Componentes instalados con Presto
emr-5.0.2	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.1	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.0.0	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Apache Spark

Apache Spark es un marco de procesamiento distribuido y un modelo de programación que lo ayuda a aplicar machine learning, procesamiento de flujos o análisis de gráficos con clústeres de Amazon EMR. Del mismo modo que Apache Hadoop, Spark es un sistema de procesamiento distribuido de código abierto utilizado frecuentemente para cargas de trabajo de big data. Sin embargo, Spark tiene varias diferencias notables con respecto a Hadoop MapReduce. Spark tiene un motor de ejecución optimizado de gráficos acíclicos dirigidos (DAG) y almacena en caché datos en memoria de forma activa, lo que puede aumentar el rendimiento especialmente para ciertos algoritmos y consultas interactivas.

Spark admite de forma nativa aplicaciones escritas en Scala, Python y Java. También incluye varias bibliotecas estrechamente integradas para SQL (Spark), aprendizaje automático (MLlib), procesamiento de flujos (Spark streaming) y procesamiento de gráficos (GraphX). Estas herramientas facilitan que Spark aproveche el marco para una amplia variedad de casos de uso.

Puede instalar Spark en un clúster de Amazon EMR junto con otras aplicaciones de Hadoop y el motor puede usar el sistema de archivos de Amazon EMR (EMRFS) para acceder directamente a los datos en Amazon S3. Hive también está integrado con Spark, por lo que puedes usar un HiveContext objeto para ejecutar scripts de Hive con Spark. Un contexto de Hive se incluye en el shell de Spark como `sqlContext`.

Para ver un ejemplo de tutorial sobre cómo configurar un clúster de EMR con Spark y analizar un conjunto de datos de muestra, consulte Tutorial: Cómo empezar con Amazon EMR en el blog de noticias. AWS

> **Important**
>
> La versión 2.3.1 de Apache Spark, disponible a partir de la versión 5.16.0 de Amazon EMR, aborda CVE-2018-8024 y CVE-2018-1334. Se recomienda migrar las versiones anteriores de Spark a la versión 2.3.1 o posteriores de Spark.

En la siguiente tabla, se muestra la versión de Spark incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Spark.

Para ver la versión de los componentes instalados con Spark en esta versión, consulte Versiones de componentes de la versión 7.9.0.

Información sobre la versión de Spark para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-7.9.0	Spark 3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

En la tabla siguiente, se muestra la versión de Spark incluida en la versión más reciente de la serie 6.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Spark.

Para ver la versión de los componentes instalados con Spark en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Spark para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.15.0	Spark 3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
		kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

> **Note**
>
> La versión 6.8.0 de Amazon EMR incluye la versión 3.3.0 de Apache Spark. Esta versión de Spark utiliza Apache Log4j 2 y el archivo `log4j2.properties` para configurar Log4j en los procesos de Spark. Si utiliza Spark en el clúster o crea clústeres de EMR con parámetros de configuración personalizados y desea actualizar a la versión 6.8.0 de Amazon EMR, debe migrar al nuevo formato de clave y clasificación de configuración `spark-log4j2` de Apache Log4j 2. Para obtener más información, consulte Migración de Apache Log4j 1.x a Log4j 2.x.

En la tabla siguiente, se muestra la versión de Spark incluida en la versión más reciente de la serie 5.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Spark.

Para ver la versión de los componentes instalados con Spark en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Spark para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.36.2	Spark 2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Temas

- Creación de un clúster con Apache Spark
- Ejecución de aplicaciones de Spark con Docker en Amazon EMR 6.x
- Usa el catálogo de AWS Glue Data Catalog con Spark en Amazon EMR
- Trabajar con una jerarquía de varios catálogos en AWS Glue Data Catalog with Spark en Amazon EMR
- Configurar Spark
- Optimización del desempeño de Spark
- Almacenamiento en caché de fragmentos de resultados de Spark
- Uso de Nvidia RAPIDS Accelerator de Nvidia para Spark
- Acceso al intérprete de comandos de Spark
- Utilice Amazon SageMaker Spark para el aprendizaje automático
- Escritura de una aplicación de Spark
- Mejora del rendimiento de Spark con Amazon S3
- Adición de un paso de Spark
- Visualización del historial de aplicaciones de Spark
- Accede a la web de Spark UIs
- Uso del conector Amazon Kinesis Data Streams de streaming estructurado de Spark
- Uso de la integración de Amazon Redshift para Apache Spark con Amazon EMR

- Historial de lanzamientos de Spark

Creación de un clúster con Apache Spark

El siguiente procedimiento crea un clúster con Spark instalado a través de las Opciones rápidas en la consola de Amazon EMR.

Como alternativa, puede usar Opciones avanzadas para personalizar aún más la configuración del clúster o para enviar pasos a fin de instalar aplicaciones mediante programación y, a continuación, ejecutar las aplicaciones personalizadas. Con alguna de las opciones de creación de clústeres, puede elegir utilizar AWS como metaalmacén de Spark SQL. Para obtener más información, consulte Usa el catálogo de AWS Glue Data Catalog con Spark en Amazon EMR.

Para lanzar un clúster con Spark instalado

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. Elija Crear clúster para utilizar Opciones rápidas.
3. Ingrese un nombre de clúster. El nombre del clúster no puede contener los caracteres <, >, $, | o ` (acento grave).
4. En Configuración de software, elija una opción de Versión.
5. En Aplicaciones, elija el paquete de aplicaciones de Spark.
6. Seleccione las demás opciones que necesite y, a continuación, elija Create cluster (Crear clúster).

> **Note**
>
> Para configurar Spark cuando se crea el clúster, consulte Configurar Spark.

Para lanzar un clúster con Spark instalado, utilice el AWS CLI

- Cree el clúster con el siguiente comando:

```
aws emr create-cluster --name "Spark cluster" --release-label emr-7.9.0 --
applications Name=Spark \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 --use-
default-roles
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Para lanzar un clúster con Spark instalado con el SDK para Java

Especifique Spark como una aplicación con `SupportedProductConfig` utilizado en `RunJobFlowRequest`.

- El siguiente ejemplo muestra cómo crear un clúster con Spark mediante Java.

```
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;
import com.amazonaws.services.elasticmapreduce.util.StepFactory;

public class Main {

        public static void main(String[] args) {
                AWSCredentials credentials_profile = null;
                try {
                        credentials_profile = new
 ProfileCredentialsProvider("default").getCredentials();
                } catch (Exception e) {
                        throw new AmazonClientException(
                                        "Cannot load credentials from .aws/
credentials file. " +
                                                        "Make sure that the
 credentials file exists and the profile name is specified within it.",
                                        e);
                }

                AmazonElasticMapReduce emr =
 AmazonElasticMapReduceClientBuilder.standard()
```

```
                                .withCredentials(new
 AWSStaticCredentialsProvider(credentials_profile))
                                .withRegion(Regions.US_WEST_1)
                                .build();

                // create a step to enable debugging in the AWS Management Console
                StepFactory stepFactory = new StepFactory();
                StepConfig enabledebugging = new StepConfig()
                                .withName("Enable debugging")
                                .withActionOnFailure("TERMINATE_JOB_FLOW")

 .withHadoopJarStep(stepFactory.newEnableDebuggingStep());

                Application spark = new Application().withName("Spark");

                RunJobFlowRequest request = new RunJobFlowRequest()
                                .withName("Spark Cluster")
                                .withReleaseLabel("emr-5.20.0")
                                .withSteps(enabledebugging)
                                .withApplications(spark)
                                .withLogUri("s3://path/to/my/logs/")
                                .withServiceRole("EMR_DefaultRole")
                                .withJobFlowRole("EMR_EC2_DefaultRole")
                                .withInstances(new JobFlowInstancesConfig()
                                                .withEc2SubnetId("subnet-12ab3c45")
                                                .withEc2KeyName("myEc2Key")
                                                .withInstanceCount(3)

 .withKeepJobFlowAliveWhenNoSteps(true)
                                                .withMasterInstanceType("m4.large")

 .withSlaveInstanceType("m4.large"));
                RunJobFlowResult result = emr.runJobFlow(request);
                System.out.println("The cluster ID is " + result.toString());
        }
}
```

Ejecución de aplicaciones de Spark con Docker en Amazon EMR 6.x

Con Amazon EMR 6.0.0, las aplicaciones de Spark pueden usar contenedores de Docker para definir las dependencias de sus bibliotecas, en lugar de instalarlas en las instancias individuales de Amazon EC2 del clúster. Para ejecutar Spark con Docker, primero debe configurar el registro de Docker y definir parámetros adicionales al enviar una aplicación Spark. Para obtener más información, consulte Configuración de la integración de Docker.

Cuando se envía la aplicación, YARN invoca Docker para extraer la imagen de Docker especificada y ejecutar la aplicación Spark dentro de un contenedor de Docker. Esto le permite definir y aislar fácilmente las dependencias. Reduce la duración del proceso de arranque o preparación de las instancias en el clúster de Amazon EMR con las bibliotecas necesarias para la ejecución del trabajo.

Consideraciones al ejecutar Spark con Docker

Cuando ejecute Spark con Docker, asegúrese de que se cumplen los siguientes requisitos previos:

- El paquete `docker` y la CLI solo deben estar instalados en los nodos principales y de tareas.
- En la versión 6.1.0 y posteriores de Amazon EMR, también puede instalar Docker en un nodo principal con los siguientes comandos.
  - ```
    sudo yum install -y docker
    sudo systemctl start docker
    ```
- El comando `spark-submit` siempre debe ejecutarse desde una instancia principal en el clúster de Amazon EMR.
- Los registros de Docker utilizados con el fin de resolver imágenes de Docker deben definirse mediante la API de clasificación con la clave de clasificación `container-executor` para definir parámetros adicionales al iniciar el clúster:
  - `docker.trusted.registries`
  - `docker.privileged-containers.registries`
- Para ejecutar una aplicación Spark en un contenedor de Docker, son necesarias las siguientes opciones de configuración:
  - `YARN_CONTAINER_RUNTIME_TYPE=docker`
  - `YARN_CONTAINER_RUNTIME_DOCKER_IMAGE={DOCKER_IMAGE_NAME}`
    ```

- Cuando utilice Amazon ECR para recuperar imágenes de Docker, debe configurar el clúster de tal forma que se autentique por sí solo. Para ello, debe utilizar la siguiente opción de configuración:
  - YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON
- En la versión 6.1.0 y posteriores de Amazon EMR, no es necesario utilizar el comando indicado `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON_HDFS` cuando la característica de autenticación automática de ECR está habilitada.
- Cualquier imagen de Docker utilizada con Spark debe tener Java instalado en la imagen de Docker.

Para obtener más información acerca de los requisitos previos, consulte Configurar la integración de Docker.

Creación de una imagen de Docker

Las imágenes de Docker se crean utilizando un archivo Dockerfile, que define los paquetes y la configuración que se van a incluir en la imagen. Los siguientes dos ejemplos de Dockerfiles utilizan y PySpark SparkR.

PySpark Dockerfile

Las imágenes de Docker creadas a partir de este Dockerfile incluyen Python 3 y el paquete Python NumPy . Este archivo Dockerfile utiliza Amazon Linux 2 y Amazon Corretto JDK 8.

```
FROM amazoncorretto:8

RUN yum -y update
RUN yum -y install yum-utils
RUN yum -y groupinstall development

RUN yum list python3*
RUN yum -y install python3 python3-dev python3-pip python3-virtualenv

RUN python -V
RUN python3 -V

ENV PYSPARK_DRIVER_PYTHON python3
ENV PYSPARK_PYTHON python3

RUN pip3 install --upgrade pip
```

```
RUN pip3 install numpy pandas

RUN python3 -c "import numpy as np"
```

Archivo Dockerfile de SparkR

Las imágenes de Docker creadas a partir de este archivo Dockerfile incluyen R y el paquete randomForest CRAN. Este archivo Dockerfile incluye Amazon Linux 2 y Amazon Corretto JDK 8.

```
FROM amazoncorretto:8

RUN java -version

RUN yum -y update
RUN amazon-linux-extras install R4

RUN yum -y install curl hostname

#setup R configs
RUN echo "r <- getOption('repos'); r['CRAN'] <- 'http://cran.us.r-project.org';
 options(repos = r);" > ~/.Rprofile

RUN Rscript -e "install.packages('randomForest')"
```

Para obtener más información sobre la sintaxis de los archivos Dockerfile, consulte la documentación de referencia de Dockerfile.

Uso de imágenes de Docker de Amazon ECR

Amazon Elastic Container Registry (Amazon ECR) es un registro de contenedores de Docker completamente administrado que facilita el almacenamiento, la administración y la implementación de imágenes de contenedores de Docker. Cuando se utiliza Amazon ECR, el clúster debe configurarse de tal forma que confíe en su instancia de ECR. Además, debe configurar la autenticación para que el clúster utilice imágenes de Docker desde Amazon ECR. Para obtener más información, consulte Configuración de YARN para acceder a Amazon ECR.

Para garantizar que los hosts de Amazon EMR puedan acceder a las imágenes almacenadas en Amazon ECR, su clúster debe tener los permisos de la política `AmazonEC2ContainerRegistryReadOnly` asociados al perfil de instancia. Para obtener más información, consulte la política `AmazonEC2ContainerRegistryReadOnly`.

En este ejemplo, el clúster debe crearse con la siguiente configuración adicional para asegurarse de que el registro de Amazon ECR sea de confianza. Sustituya el *123456789123.dkr.ecr.us-east-1.amazonaws.com* punto de conexión por su punto de conexión de Amazon ECR.

```
[
  {
    "Classification": "container-executor",
    "Configurations": [
      {
        "Classification": "docker",
        "Properties": {
          "docker.privileged-containers.registries":
 "local,centos,123456789123.dkr.ecr.us-east-1.amazonaws.com",
          "docker.trusted.registries": "local,centos,123456789123.dkr.ecr.us-
east-1.amazonaws.com"
        }
      }
    ],
    "Properties": {}
  }
]
```

Uso PySpark con Amazon ECR

El siguiente ejemplo usa el PySpark Dockerfile, que se etiquetará y se cargará en Amazon ECR. Tras cargar el Dockerfile, puede ejecutar el PySpark trabajo y consultar la imagen de Docker de Amazon ECR.

Tras lanzar el clúster, utilice SSH para conectarse a un nodo principal y ejecute los siguientes comandos para crear la imagen de Docker local a partir del ejemplo de Dockerfile. PySpark

Primero, cree un directorio y un archivo Dockerfile.

```
mkdir pyspark
vi pyspark/Dockerfile
```

Pegue el contenido del PySpark Dockerfile y ejecute los siguientes comandos para crear una imagen de Docker.

```
sudo docker build -t local/pyspark-example pyspark/
```

Cree el repositorio `emr-docker-examples` de ECR para los ejemplos.

```
aws ecr create-repository --repository-name emr-docker-examples
```

Etiquete y cargue la imagen creada localmente en ECR y *123456789123.dkr.ecr.us-east-1.amazonaws.com* sustitúyala por su terminal de ECR.

```
sudo docker tag local/pyspark-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-
docker-examples:pyspark-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
```

Utilice SSH para conectarse al nodo principal y preparar un script de Python con el nombre de archivo `main.py`. Pegue el siguiente contenido en el archivo `main.py` y guárdelo.

```
from pyspark.sql import SparkSession
spark = SparkSession.builder.appName("docker-numpy").getOrCreate()
sc = spark.sparkContext

import numpy as np
a = np.arange(15).reshape(3, 5)
print(a)
```

En Amazon EMR 6.0.0, para enviar el trabajo, haga referencia al nombre de la imagen de Docker. Defina los parámetros de configuración adicionales para asegurarse de que la ejecución del trabajo utiliza Docker como tiempo de ejecución. Cuando se utiliza Amazon ECR, `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` debe hacer referencia al archivo `config.json` que contiene las credenciales utilizadas para autenticarse en Amazon ECR.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
```

```
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--num-executors 2 \
main.py -v
```

En la versión 6.1.0 y posteriores de Amazon EMR, para enviar el trabajo, haga referencia al nombre de la imagen de Docker. Cuando la autenticación automática de ECR esté habilitada, ejecute el siguiente comando.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--num-executors 2 \
main.py -v
```

Cuando se complete el trabajo, anote el ID de la aplicación YARN y utilice el siguiente comando para obtener el resultado del PySpark trabajo.

```
yarn logs --applicationId application_id | grep -C2 '\[\['
LogLength:55
LogContents:
[[ 0  1  2  3  4]
 [ 5  6  7  8  9]
 [10 11 12 13 14]]
```

Uso de SparkR con Amazon ECR

En el ejemplo siguiente se utiliza el archivo Dockerfile de SparkR, que se etiquetará y se cargará en ECR. Una vez cargado el Dockerfile, puede ejecutar el trabajo de SparkR y hacer referencia a la imagen de Docker desde Amazon ECR.

Después de iniciar el clúster, utilice SSH para conectarse a un nodo principal y ejecute los siguientes comandos para crear la imagen de Docker local desde el ejemplo de archivo Dockerfile de SparkR.

Primero, cree un directorio y el archivo Dockerfile.

```
mkdir sparkr
```

```
vi sparkr/Dockerfile
```

Pegue el contenido del archivo Dockerfile de SparkR y ejecute los siguientes comandos para crear una imagen de Docker.

```
sudo docker build -t local/sparkr-example sparkr/
```

Etiquete y cargue la imagen creada localmente en Amazon ECR y sustitúyala por su *123456789123.dkr.ecr.us-east-1.amazonaws.com* punto de conexión de Amazon ECR.

```
sudo docker tag local/sparkr-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-
docker-examples:sparkr-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
```

Utilice SSH para conectarse al nodo principal y preparar un script de R con el nombre `sparkR.R`. Pegue los siguientes contenidos en el archivo `sparkR.R`:

```
library(SparkR)
sparkR.session(appName = "R with Spark example", sparkConfig =
 list(spark.some.config.option = "some-value"))

sqlContext <- sparkRSQL.init(spark.sparkContext)
library(randomForest)
# check release notes of randomForest
rfNews()

sparkR.session.stop()
```

En Amazon EMR 6.0.0, para enviar el trabajo, haga referencia al nombre de la imagen de Docker. Defina los parámetros de configuración adicionales para asegurarse de que la ejecución del trabajo utiliza Docker como tiempo de ejecución. Cuando se utiliza Amazon ECR, `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` debe hacer referencia al archivo `config.json` que contiene las credenciales utilizadas para autenticarse en ECR.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
```

```
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
sparkR.R
```

En la versión 6.1.0 y posteriores de Amazon EMR, para enviar el trabajo, haga referencia al nombre de la imagen de Docker. Cuando la autenticación automática de ECR esté habilitada, ejecute el siguiente comando.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
sparkR.R
```

Cuando el trabajo se haya completado, tome nota del ID de aplicación de YARN y utilice el siguiente comando para obtener la salida del trabajo de SparkR. Este ejemplo incluye pruebas para asegurarse de que la biblioteca randomForest, la versión instalada y las notas de la versión están disponibles.

```
yarn logs --applicationId application_id | grep -B4 -A10 "Type rfNews"
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.
Wishlist (formerly TODO):

* Implement the new scheme of handling classwt in classification.

* Use more compact storage of proximity matrix.

* Allow case weights by using the weights in sampling?

=============================================================================
Changes in 4.6-14:
```

Usa el catálogo de AWS Glue Data Catalog con Spark en Amazon EMR

Con la versión 5.8.0 o posterior de Amazon EMR, puedes configurar Spark para que utilice el catálogo de datos de AWS Glue como metaalmacén de Apache Hive. Recomendamos esta configuración cuando necesites un metastore de Hive persistente o uno de Hive compartido por distintos clústeres, servicios, aplicaciones o cuentas. AWS

Con Amazon EMR versión 6.5.0 o posterior, puedes configurar Spark para que utilice el catálogo de datos de AWS Glue con Apache Iceberg.

Con la versión 7.5.0 o posterior de Amazon EMR, puedes configurar Spark para que utilice el catálogo de datos de AWS Glue como su catálogo REST de Iceberg.

AWS Glue es un servicio de extracción, transformación y carga (ETL) totalmente gestionado que permite clasificar los datos, limpiarlos, enriquecerlos y moverlos de forma fiable entre varios almacenes de datos de forma sencilla y rentable. El catálogo de datos de AWS Glue proporciona un repositorio de metadatos unificado en una variedad de fuentes y formatos de datos, y se integra con Amazon EMR, así como con Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena y cualquier aplicación compatible con el metaalmacén de Apache Hive. AWS Los rastreadores de Glue pueden deducir automáticamente el esquema a partir de los datos de origen en Amazon S3 y almacenar los metadatos asociados en el catálogo de datos. Para obtener más información sobre el catálogo de datos, consulte Cómo rellenar el catálogo de datos de AWS Glue en la Guía del desarrollador de AWS Glue.

Se aplican cargos separados por AWS Glue. Hay una tarifa mensual para almacenar y acceder a los metadatos del catálogo de datos, una tarifa por hora que se factura por minuto para los trabajos de ETL de AWS Glue y el tiempo de ejecución del rastreador, y una tarifa por hora que se factura por minuto para cada terminal de desarrollo aprovisionado. El Catálogo de datos le permite almacenar hasta un millón de objetos sin cargo alguno. Si almacena más de un millón de objetos, se le cobrará 1 USD por cada 100 000 objetos que sobrepasen esa cifra. En el Catálogo de datos, se consideran objetos las tablas, las particiones y las bases de datos. Para obtener más información, consulte Precios de Glue.

> **Important**
>
> Si creó tablas con Amazon Athena o Amazon Redshift Spectrum antes del 14 de agosto de 2017, las bases de datos y las tablas se almacenan en un catálogo gestionado por Athena,

que es independiente del catálogo de datos de Glue. AWS Para integrar Amazon EMR con estas tablas, debe actualizar al catálogo de datos de AWS Glue. Para obtener más información, consulte Actualización al catálogo de datos de AWS Glue en la Guía del usuario de Amazon Athena.

Especificar AWS Glue Data Catalog como metastore de Apache Hive

Puedes especificar el catálogo de datos de AWS Glue como metastore mediante la AWS Management Console API AWS CLI, o Amazon EMR. Cuando utiliza la CLI o la API, utiliza la clasificación de configuración de Spark para especificar el catálogo de datos. Además, con Amazon EMR 5.16.0 y versiones posteriores, puede usar la clasificación de configuración para especificar un catálogo de datos en otro. Cuenta de AWS Cuando utiliza la consola, puede especificar el Catálogo de datos mediante Opciones avanzadas u Opciones rápidas.

> **Note**
>
> La opción de usar AWS Glue Data Catalog también está disponible con Zeppelin porque Zeppelin se instala con los componentes de Spark.

Console

Para especificar AWS Glue Data Catalog como el metaalmacén de Apache Hive con la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En Amazon EMR, EC2 en el panel de navegación izquierdo, selecciona Clústeres y, a continuación, selecciona Crear clúster.
3. En Paquete de aplicaciones, seleccione Spark o Personalizado. Si personaliza el clúster, asegúrese de seleccionar Zeppelin o Spark como una de sus aplicaciones.
4. En Configuración del Catálogo de datos de AWS Glue, seleccione la casilla Usar para metadatos de la tabla de Spark.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para especificar el catálogo de datos de AWS Glue como el metaalmacén de Apache Hive con el AWS CLI

Para obtener más información sobre cómo especificar una clasificación de configuración mediante la AWS CLI API de Amazon EMR, consulte. Configuración de aplicaciones

- Especifique el valor de `hive.metastore.client.factory.class` utilizando la clasificación `spark-hive-site` que se muestra en el ejemplo siguiente:

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

Para especificar un catálogo de datos en una AWS cuenta diferente, añada la `hive.metastore.glue.catalogid` propiedad tal y como se muestra en el siguiente ejemplo. Sustituya *`acct-id`* por la cuenta de AWS del Catálogo de datos.

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Especificar AWS Glue Data Catalog como catálogo de Apache Iceberg

Puedes especificar el catálogo de datos de AWS Glue como la implementación del catálogo de Apache Iceberg o el punto final del catálogo REST de Apache Iceberg, mediante la AWS Management Console API Amazon EMR o en la AWS CLI configuración de tiempo de ejecución de sesión de Spark. Cuando utiliza la CLI o la API, utiliza la clasificación de configuración de Spark para especificar el catálogo de datos. Para obtener más información, consulte Especificar el catálogo de datos de AWS Glue como el catálogo de Apache Iceberg.

Permisos de IAM

El perfil de EC2 instancia de un clúster debe tener permisos de IAM para las acciones de AWS Glue. Además, si habilita el cifrado de los objetos del catálogo de datos de AWS Glue, el rol también debe poder cifrar, descifrar y generar lo que AWS KMS key se utiliza para el cifrado.

Permisos para las acciones de AWS Glue

Si utiliza el perfil de EC2 instancia predeterminado para Amazon EMR, no es necesario realizar ninguna acción. La política `AmazonElasticMapReduceforEC2Role` gestionada que se adjunta a la `EMR_EC2_DefaultRole` permite todas las acciones de AWS Glue necesarias. Sin embargo, si especificas un perfil de EC2 instancia y permisos personalizados, debes configurar las acciones de AWS Glue adecuadas. Utilice la política administrada `AmazonElasticMapReduceforEC2Role` como punto de partida. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia) en la Guía de administración de Amazon EMR.

Permisos para cifrar y descifrar AWS Glue Data Catalog

El perfil de instancia necesita permiso para cifrar y descifrar los datos con su clave. No es necesario configurar estos permisos si se cumplen las dos instrucciones siguientes:

- Para activar el cifrado de los objetos del catálogo de datos de AWS Glue, utilice claves gestionadas para AWS Glue.
- Usas un clúster que está en el mismo lugar Cuenta de AWS que el catálogo de datos de AWS Glue.

De lo contrario, debes añadir la siguiente declaración a la política de permisos adjunta a tu perfil de EC2 instancia.

```
[
```

```
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Para obtener más información sobre el cifrado del catálogo de datos de AWS Glue, consulte Cómo cifrar el catálogo de datos en la Guía para desarrolladores de AWS Glue.

Permisos basados en recursos

Si usa AWS Glue junto con Hive, Spark o Presto en Amazon EMR, AWS Glue admite políticas basadas en recursos para controlar el acceso a los recursos del catálogo de datos. Estos recursos incluyen bases de datos, tablas, conexiones y funciones definidas por el usuario. Para obtener más información, consulte Políticas de recursos de AWS Glue en la Guía para desarrolladores de AWS Glue.

Al utilizar políticas basadas en recursos para limitar el acceso a AWS Glue desde Amazon EMR, el principal que especifique en la política de permisos debe ser el ARN de rol asociado al perfil de EC2 instancia que se especifica al crear un clúster. Por ejemplo, para una política basada en recursos adjunta a un catálogo, puede especificar el ARN del rol para el rol de servicio predeterminado para las EC2 instancias de clúster, *EMR_EC2_DefaultRole* como el`Principal`, con el formato que se muestra en el siguiente ejemplo:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*Puede ser diferente del ID de la cuenta de AWS Glue. Esto permite el acceso desde clústeres de EMR en diferentes cuentas. Puede especificar varias entidades principales, cada una de ellas desde una cuenta diferente.

Consideraciones a la hora de utilizar el Catálogo de datos de AWS Glue

Ten en cuenta lo siguiente cuando utilices AWS Glue Data Catalog como un metaalmacén de Apache Hive con Spark:

- Una base de datos predeterminada sin un URI de ubicación genera errores al crear una tabla. Para solucionar este problema, utilice la cláusula `LOCATION` para especificar la ubicación de un bucket, como `s3://amzn-s3-demo-bucket1`, cuando utilice `CREATE TABLE`. Otra opción consiste en crear tablas dentro de una base de datos distinta de la base de datos predeterminada.
- No se admite el cambio de nombre de las tablas desde AWS Glue.
- Al crear una tabla de Hive sin especificar un valor de `LOCATION`, los datos de la tabla se almacenan en la ubicación especificada por la propiedad `hive.metastore.warehouse.dir`. De manera predeterminada, se trata de una ubicación en HDFS. Si otro clúster necesita acceder a la tabla, se produce un error a menos que tenga los permisos adecuados para el clúster que creó la tabla. Además, dado que el almacenamiento en HDFS es transitorio, si el clúster termina, se pierden los datos de la tabla y, por lo tanto, será necesario volver a crearla. Le recomendamos que especifique una `LOCATION` en Amazon S3 cuando cree una tabla Hive con AWS Glue. Como alternativa, puede utilizar la clasificación de configuración `hive-site` para especificar una ubicación en Amazon S3 para `hive.metastore.warehouse.dir`, que se aplica a todas las tablas de Hive. Si se crea una tabla en una ubicación de HDFS y el clúster que la creó sigue ejecutándose, puedes actualizar la ubicación de la tabla a Amazon S3 desde AWS Glue. Para obtener más información, consulte Cómo trabajar con tablas en la consola de AWS Glue en la Guía para desarrolladores de AWS Glue.
- No se admiten valores de partición que contengan comillas y apóstrofes, por ejemplo, `PARTITION (owner="Doe's").`
- Las estadísticas de columnas son compatibles con emr-5.31.0 y versiones posteriores.
- No se admite el uso de la autorización de Hive. Como alternativa, considere la posibilidad de utilizar políticas basadas en recursos de AWS Glue. Para obtener más información, consulte Uso de políticas basadas en recursos para Amazon EMR Access to Glue AWS Data Catalog.

Tenga en cuenta lo siguiente cuando utilice AWS Glue Data Catalog como catálogo REST de Apache Iceberg con Spark:

- Si usa el catálogo de sesiones de Spark con Iceberg, como se describe enDiferencias de configuración cuando se usa Iceberg SparkCatalog y SparkSessionCatalog, debe configurar el

catálogo de datos de AWS Glue como el metaalmacén de Apache Hive, además de configurar el catálogo de datos de AWS Glue como el catálogo REST de Apache Iceberg.

- El punto final IRC de AWS Glue Data Catalog solo admite el esquema de autenticación Amazon SigV4. OAuth no es compatible. Para OAuth los usuarios, utilice el Centro de identidad de IAM para configurar el acceso. Consulte Connecting Lake Formation con IAM Identity Center.
- El catálogo REST de AWS Glue Iceberg no admite todas las operaciones en código abierto.

Trabajar con una jerarquía de varios catálogos en AWS Glue Data Catalog with Spark en Amazon EMR

Puede registrar su clúster de Amazon EMR para acceder al catálogo de datos de AWS Glue, que pone las tablas y otros recursos del catálogo a disposición de varios consumidores. AWS Glue Data Catalog admite una jerarquía de varios catálogos, que unifica los datos en los lagos de datos de Amazon S3. También proporciona una API de metastore de Hive y una API REST Apache Iceberg de código abierto para acceder a los datos. Estas funciones están disponibles para Amazon EMR y otros servicios, como Amazon Athena Amazon Redshift.

Cómo se organizan los recursos del catálogo

Al crear recursos en el catálogo de datos de AWS Glue, puede acceder a ellos desde cualquier motor SQL que admita la API REST de Apache Iceberg o el metaalmacén de Hive. AWS Lake Formation gestiona los permisos.

En AWS Glue Data Catalog, los datos se organizan en una jerarquía lógica de catálogos, bases de datos y tablas:

- Catálogo: contenedor lógico que contiene objetos de un banco de datos, como esquemas o tablas.
- Catálogo para almacenar tablas de Redshift Managed Storage (RMS): cuando administra catálogos para almacenar tablas RMS, puede acceder a estas tablas mediante Iceberg.
- Base de datos: organiza los objetos de datos, como tablas y vistas, en un catálogo.
- Tablas y vistas: objetos de datos de una base de datos que proporcionan una capa de abstracción con un esquema comprensible. Proporcionan una capa para acceder a los datos subyacentes, que pueden estar en varios formatos y en diferentes ubicaciones.

Configuración de un catálogo de datos para su uso con Amazon EMR

Para empezar, debe configurar el catálogo para que sea compatible con las herramientas de Amazon EMR. El catálogo de datos de AWS Glue es compatible con Hive Metastore y con Iceberg REST. APIs

Configuración de Amazon EMR con un metastore de Hive

Para obtener información sobre cómo configurarlo, consulta la compatibilidad del catálogo de datos de AWS Glue para trabajos de Spark en la Guía del usuario de AWS Glue. En este tema se describe cómo configurar el catálogo de datos de AWS Glue como un metaalmacén de Hive y hacer que esté disponible como punto final. Además, hay documentación de Amazon EMR disponible que muestra cómo especificar el catálogo de datos de AWS Glue como un metaalmacén de Spark, en Use the AWS Glue Data Catalog como metastore de Apache Hive para Spark.

Permisos para acceder a los recursos del catálogo de datos de AWS Glue

En esta sección se describen los requisitos de la política de IAM para utilizar las herramientas de Amazon EMR con los datos del catálogo. Tras registrar el clúster en el catálogo de datos de AWS Glue, necesitará los siguientes permisos para descubrir la creación y los cambios en el catálogo de datos creado posteriormente:

- pegamento: GetCatalog
- pegamento: GetCatalogs
- conjuntos: AssumeRole
- conjuntos: TagSession
- conjuntos: SetContext
- conjuntos: SetSourceIdentity

En la mayoría de los casos, al asignar permisos, se recomienda crear un rol de IAM y asignarle permisos.

Además, para consultar los datos del catálogo, debe configurar los permisos para el catálogo de datos mediante. AWS Lake Formation Para obtener más información sobre cómo configurar los permisos para los catálogos de datos AWS Lake Formation, consulte Concesión y revocación de permisos en los recursos del catálogo de datos.

Tras crear y configurar el clúster y establecer los permisos en los objetos del catálogo, puede enviar trabajos para consultar y procesar datos.

Configurar Spark para acceder a una jerarquía de varios catálogos en AWS Glue Data Catalog

Con EMR 7.5, puedes configurar Spark para que utilice la jerarquía multicálogos de AWS Glue. Una jerarquía de varios catálogos te permite:

- Lleve sus datos de Redshift Managed Storage (RMS), como tablas, vistas y vistas materializadas de los almacenes de datos de Amazon Redshift existentes, a Glue Data Catalog. AWS Puede consultar estos objetos mediante EMR on y EMR EC2 Serverless.
- Cree catálogos RMS, AWS Glue Data Catalog y almacene datos en RMS con ZeroETL y consulte los datos con motores de consulta compatibles con Iceberg.
- Cree tablas Iceberg gestionadas en AWS Glue Data Catalog con una gestión de almacenamiento completa que incluye compactación, instantáneas y retención.

Conexión a un catálogo múltiple al inicializar una sesión de Spark

Los siguientes ejemplos muestran cómo usar Spark shell interactivo, Spark submit o Amazon EMR Notebooks para trabajar AWS con la jerarquía de múltiples catálogos de Glue.

spark-shell

1. Conéctese al nodo principal utilizando SSH. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.
2. Introduzca el siguiente comando para iniciar el shell de Spark. Para usar la PySpark carcasa, sustitúyala por. `spark-shell` `pyspark`

```
spark-shell \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. Conéctese al nodo principal utilizando SSH. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.
2. Ingresa el siguiente comando para iniciar la sesión de Spark para Spark.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

Para iniciar una sesión de Spark con los Cuadernos de Amazon EMR, configure su sesión de Spark con el comando mágico `%%configure` de su cuaderno de Amazon EMR, como en el siguiente ejemplo. Para obtener más información, consulte Usar la magia de Cuadernos de Amazon EMR en la Guía de administración de Amazon EMR.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
}
```

CLI

Para inicializar una sesión de Spark mediante la CLI, ejecuta el siguiente ejemplo. Para obtener más información sobre cómo especificar una clasificación de configuración mediante la AWS CLI API de Amazon EMR, consulte Configurar aplicaciones.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

Inicie una sesión de Spark en Redshift Managed Storage con AWS Glue Data Catalog

El siguiente comando de ejemplo inicializa una sesión de Spark con el catálogo de datos de AWS Glue.

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
  --conf spark.sql.catalog.rms.type=glue \
  --conf spark.sql.catalog.rms.glue.id=Glue RMS catalog ID \
  --conf spark.sql.defaultCatalog=rms
  --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

El siguiente ejemplo inicializa una sesión de Spark con la API REST de Iceberg y Redshift Managed Storage with Glue AWS Data Catalog.

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
  --conf spark.sql.catalog.rms.type=rest \
  --conf spark.sql.catalog.rms.warehouse=glue RMS catalog ID \
  --conf spark.sql.catalog.rms.uri=Glue endpoint URI/iceberg \
  --conf spark.sql.catalog.rms.rest.sigv4-enabled=true \
  --conf spark.sql.catalog.rms.rest.signing-name=glue \
  --conf spark.sql.defaultCatalog=rms \
  --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Para obtener más información sobre el uso de una jerarquía de catálogos múltiples de AWS Glue con Spark Iceberg, consulta Usar un clúster de Iceberg con Spark.

Consideraciones y limitaciones de una configuración de varios catálogos

- No se admite el uso de una jerarquía de varios catálogos con el metaalmacén de Apache Hive.
- El uso de una jerarquía de varios catálogos con Apache Iceberg no permite recurrir al metaalmacén de Apache Hive cuando se utiliza. `SparkSessionCatalog`
- Los EMR en EC2 los clústeres con la función Runtime no admiten la jerarquía de varios catálogos.
- Los EMR en EC2 los clústeres habilitados AWS Lake Formation no admiten la jerarquía de varios catálogos.

Configurar Spark

Puede configurar Spark en Amazon EMR con clasificaciones de configuración. Para obtener más información sobre cómo las clasificaciones de configuración, consulte Configuración de aplicaciones.

Entre las clasificaciones de configuración para Spark en Amazon EMR se incluyen los siguientes:

- **`spark`**: establece la propiedad `maximizeResourceAllocation` en true o false. Si es true, Amazon EMR configura automáticamente las propiedades `spark-defaults` en función de la configuración de hardware del clúster. Para obtener más información, consulte Uso de `maximizeResourceAllocation`.
- **`spark-defaults`**: establece los valores del archivo `spark-defaults.conf`. Para obtener más información, consulte Spark configuration en la documentación de Spark.
- **`spark-env`**: establece los valores del archivo `spark-env.sh`. Para obtener más información, consulte Environment variables en la documentación de Spark.
- **`spark-hive-site`**: establece los valores del archivo `hive-site.xml` para Spark.
- **`spark-log4j`**: (versión 6.7.x y anteriores de Amazon EMR) establece los valores del archivo `log4j.properties`. Para obtener más información, consulte el archivo log4j.properties.template en GitHub.
- **`spark-log4j2`**: (versión 6.8.0 y posteriores de Amazon EMR) establece los valores del archivo `log4j2.properties`. Para obtener más información, consulte el archivo log4j2.properties.template en GitHub.

- **`spark-metrics`**: establece los valores del archivo `metrics.properties`. Para conocer la configuración y obtener más información, consulte el archivo metrics.properties.template en GitHub y Metrics en la documentación de Spark.

> **Note**
>
> Si va a migrar cargas de trabajo de Spark a Amazon EMR desde otra plataforma, le recomendamos que pruebe las cargas de trabajo con los Valores predeterminados de Spark establecidos por Amazon EMR antes de agregar configuraciones personalizadas. La mayoría de los clientes observan mejoras de rendimiento con nuestra configuración predeterminada.

Temas

- Valores predeterminados de Spark establecidos por Amazon EMR
- Configuración de la recopilación de elementos no utilizados de Spark en Amazon EMR 6.1.0
- Uso de maximizeResourceAllocation
- Configuración del comportamiento de retirada de nodos
- Variable de entorno Spark ThriftServer
- Cambio de la configuración predeterminada de Spark
- Migración de Apache Log4j 1.x a Log4j 2.x

Valores predeterminados de Spark establecidos por Amazon EMR

En la siguiente tabla, se muestra cómo Amazon EMR establece valores los predeterminados en `spark-defaults` que afectan a las aplicaciones.

Valores predeterminados de Spark establecidos por Amazon EMR

Opción	Descripción	Valor predeterminado
`spark.executor.memory`	Cantidad de memoria que utilizar por proceso de ejecutor. Por ejemplo: `1g`, `2g`.	Esta configuración se determina según los tipos de instancias principales y de tareas del clúster.

Opción	Descripción	Valor predeterminado
`spark.executor.cores`	El número de núcleos que se va a utilizar en cada ejecutor.	Esta configuración se determina según los tipos de instancias principales y de tareas del clúster.
`spark.dynamicAllocation.enabled`	Si se establece en true, utilice la asignación de recursos dinámica para escalar y reducir verticalmente el número de ejecutores registrados en una aplicación en función de la carga de trabajo.	`true` (con la versión 4.4.0 y posteriores de Amazon EMR) **Note** Amazon EMR configura automáticamente Spark Shuffle Service.
`spark.sql.hive.advancedPartitionPredicatePushdown.enabled`	Si se establece en true, se habilita la inserción avanzada de predicados de particiones en el metaalmacén de Hive.	`true`

Opción	Descripción	Valor predeterminado
`spark.sql.hive.stringLikePartitionPredicatePushdown.enabled`	Inserta los filtros `startsWith` , `contains` y `endsWith` en el metaalmacén de Hive. **Note** Glue no admite la inserción de predicados para `startsWith` , `contains` ni `endsWith`. Si utiliza el metaalmacén de Glue y detecta errores debido a la inserción de predicados para estas funciones, establezca esta configuración en `false`.	`true`

Configuración de la recopilación de elementos no utilizados de Spark en Amazon EMR 6.1.0

Al establecer configuraciones personalizadas de recopilación de elementos no utilizados con `spark.driver.extraJavaOptions` y `spark.executor.extraJavaOptions`, se produce un error en el lanzamiento del controlador o el ejecutor con Amazon EMR 6.1 debido a un conflicto de configuración de recopilación de elementos no utilizados con Amazon EMR 6.1.0. En el caso de Amazon EMR 6.1.0, la configuración predeterminada de recopilación de elementos no utilizados se establece mediante `spark.driver.defaultJavaOptions` y `spark.executor.defaultJavaOptions`. Esta configuración se aplica únicamente a Amazon EMR 6.1.0. Las opciones de JVM que no estén relacionadas con la recopilación de elementos no utilizados, como las que permiten configurar el registro (`-verbose:class`), todavía se

pueden establecer mediante `extraJavaOptions`. Para obtener más información, consulte Spark application properties.

Uso de `maximizeResourceAllocation`

Si quiere configurar los ejecutores para utilizar el máximo de recursos posible en cada nodo de un clúster, establezca `maximizeResourceAllocation` en `true` en la clasificación de configuración `spark`. La opción `maximizeResourceAllocation` es específica de Amazon EMR. Al habilitar `maximizeResourceAllocation`, Amazon EMR calcula el máximo de recursos de computación y de memoria disponibles para un ejecutor en una instancia del grupo de instancias principales. A continuación, establece la configuración `spark-defaults` correspondiente en función de los valores máximos calculados.

Amazon EMR calcula el máximo de recursos de computación y de memoria disponibles para un ejecutor basado en un tipo de instancia de la flota de instancias principales. Dado que cada flota de instancias puede tener distintos tipos y tamaños de instancias dentro de una flota, la configuración de ejecutor que utiliza Amazon EMR puede no ser la mejor para sus clústeres, por lo que no recomendamos utilizar la configuración predeterminada cuando se utilice la asignación máxima de recursos. Configure ajustes personalizados para los clústeres de su flota de instancias.

> **Note**
>
> No debe usar la `maximizeResourceAllocation` opción en clústeres con otras aplicaciones distribuidas, por ejemplo. HBase Amazon EMR utiliza configuraciones de YARN personalizadas para las aplicaciones distribuidas, lo que puede entrar en conflicto con `maximizeResourceAllocation` y hacer que se produzcan errores en las aplicaciones de Spark.

A continuación, se muestra un ejemplo de una clasificación de configuración de Spark con la opción `maximizeResourceAllocation` establecida en `true`.

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
    }
  }
```

```
]
```

Opciones configuradas en **`spark-defaults`** cuando **`maximizeResourceAllocation`** está habilitado

Opción	Descripción	Valor
spark.default.parallelism	Transformaciones como unir y paralelizar RDDs devuelven el número predeterminado de particiones cuando el usuario no lo establece. reduceByKey	2X número de núcleos de CPU disponibles para contenedores de YARN.
spark.driver.memory	Cantidad de memoria que se utilizará para el proceso del controlador, es decir, dónde se SparkContext inicializa. (por ejemplo, 1 g, 2 g).	El ajuste se configura en función de los tipos de instancia en el clúster. Sin embargo, dado que la aplicación del controlador de Spark puede ejecutarse en la instancia principal o en una de las instancias básicas (por ejemplo, en un cliente de YARN y modos de clústeres , respectivamente), esto se establece en función del más pequeño de los tipos de instancia de estos dos grupos de instancias.
spark.executor.memory	Cantidad de memoria que se va a utilizar por cada proceso de ejecutor, por ejemplo 1g, 2g.	El ajuste se configura en función de los tipos de instancias secundarias y de tareas del clúster.
spark.executor.cores	El número de núcleos que se va a utilizar en cada ejecutor.	El ajuste se configura en función de los tipos de instancias secundarias y de tareas del clúster.

Opción	Descripción	Valor
spark.executor.instances	El número de ejecutores.	El ajuste se configura en función de los tipos de instancias secundarias y de tareas del clúster. Se define a menos que `spark.dynamicAllocation.enabled` defina explícitamente en true al mismo tiempo.

Configuración del comportamiento de retirada de nodos

Con la versión 5.9.0 o una posterior de Amazon EMR, Spark en Amazon EMR incluye un conjunto de características para ayudar a garantizar que Spark gestione correctamente la terminación de los nodos debido a una solicitud de cambio de tamaño manual o de política de escalado automático. Amazon EMR implementa un mecanismo de lista de denegación en Spark que se basa en el mecanismo de retirada de YARN. Este mecanismo ayuda a evitar que se programen tareas nuevas en un nodo que se está retirando, al tiempo que permite que finalicen las tareas que ya se están ejecutando. Además, existen características para ayudar a recuperar los trabajos de Spark más rápido si se pierden bloques de reorganización cuando termina un nodo. El proceso de recálculo se activa antes, y está optimizado para realizarse más rápido, con menos reintentos de etapas, y se puede evitar que se produzcan errores en los trabajos debido a problemas de recuperación provocados por los bloques de reorganización que faltan.

> **Important**
>
> La configuración `spark.decommissioning.timeout.threshold` se agregó en la versión 5.11.0 de Amazon EMR para mejorar la resiliencia de Spark cuando se utilizan instancias de spot. En las versiones anteriores, cuando un nodo utilizaba una instancia de spot y la instancia se terminaba debido al precio de puja, era posible que Spark no pudiera gestionar la terminación correctamente. Los tareas pueden fallar y los recálculos de reorganización pueden tardar mucho tiempo. Por este motivo, le recomendamos que utilice la versión 5.11.0 o posterior si utiliza instancias de spot.

Configuración de la retirada de nodos de Spark

Opción	Descripción	Valor predeterminado
`spark.blacklist.decommissioning.enabled`	Cuando se establece en `true`, Spark agrega a la lista de denegación los nodos que tienen el estado `decommissioning` en YARN. Spark no programa tareas nuevas en los ejecutores activos en ese nodo. A las tareas que se encuentran en ejecución se les permite completarse.	`true`
`spark.blacklist.decommissioning.timeout`	Cantidad de tiempo que permanece en la lista de denegación un nodo cuyo estado es `decommissioning`. De forma predeterminada, este valor se establece en una hora, que también es el valor predeterminado de `yarn.resourcemanager.decommissioning.timeout`. Para asegurarse de que un nodo permanece en la lista de denegación durante todo su periodo de retirada, establezca un valor mayor o igual que `yarn.resourcemanager.decommissioning.timeout`. Una vez transcurrido el tiempo de espera de desmantelamiento, el nodo pasa a un `decommissioned` estado y	`1h`

Opción	Descripción	Valor predeterminado
	Amazon EMR puede terminar la instancia del nodo. EC2 Si hay alguna tarea en ejecución cuando finaliza el tiempo de espera, se pierde o se cancela y se reprograma en los ejecutores activos de los otros nodos.	

Opción	Descripción	Valor predeterminado
`spark.decommissioning.timeout.threshold`	Disponible en la versión 5.11.0 o posteriores de Amazon EMR. Se especifica en segundos. Cuando un nodo pasa al estado Retirando, si el host lo va a retirar en un periodo de tiempo igual o menor que este valor, Amazon EMR no solo agrega el nodo a la lista de denegación, sino que también limpia el estado del host (según se especifique mediante `spark.resourceManager.cleanupExpiredHost`) sin necesidad de esperar a que el nodo cambie al estado Retirado. Esto permite que Spark gestione mejor las terminaciones de instancias de spot, ya que las instancias de spot se retiran dentro de un tiempo de espera de 20 segundos, independientemente del valor de `yarn.resourcemager.decommissioning.timeout`, que posiblemente no proporcione a los demás nodos tiempo suficiente para leer los archivos de reorganización.	`20s`

Opción	Descripción	Valor predeterminado
`spark.resourceManager.cleanupExpiredHost`	Cuando se establece en `true`, Spark anula el registro de todos los datos almacenados en la memoria caché y los bloques de reorganización que se almacenan en ejecutores en los nodos que tienen el estado `decommissioned` . Esto acelera el proceso de recuperación.	`true`
`spark.stage.attempt.ignoreOnDecommissionFetchFailure`	Cuando se establece en `true`, ayuda a evitar errores en las etapas de Spark y que al final el trabajo no se complete debido a un número excesivo de recuperaciones desde los nodos retirados. Los errores en las recuperaciones de los bloques de reorganización de un nodo con el estado `decommissioned` no cuentan para el cálculo del número máximo de errores de recuperación consecutivos.	`true`

Variable de entorno Spark ThriftServer

Spark establece la variable de entorno Hive Thrift Server Port, `HIVE_SERVER2_THRIFT_PORT`, en 10 001.

Cambio de la configuración predeterminada de Spark

Puede cambiar la configuración predeterminada en `spark-defaults.conf` utilizando la clasificación de configuración `spark-defaults` al crear el clúster o el ajuste `maximizeResourceAllocation` en la clasificación de configuración `spark`.

Los siguientes procedimientos muestran cómo modificar los ajustes mediante la CLI o la consola.

Para crear un clúster con el valor de spark.executor.memory definido en 2g mediante la CLI

- Cree un clúster con Spark instalado y el valor de `spark.executor.memory` establecido en 2g con el siguiente comando, que hace referencia a un archivo (`myConfig.json`) almacenado en Amazon S3.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

`myConfig.json`:

```
[
    {
      "Classification": "spark-defaults",
      "Properties": {
        "spark.executor.memory": "2G"
      }
    }
  ]
```

Para crear un clúster con el valor de spark.executor.memory definido en 2g mediante la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. Elija Spark.
4. En Edit software settings (Editar configuración de software), deje seleccionada la opción Enter configuration (Escribir la configuración) y especifique la siguiente configuración:

```
classification=spark-defaults,properties=[spark.executor.memory=2G]
```

5. Seleccione otras opciones, elija y, a continuación, elija Create cluster (Crear clúster).

Para configurar maximizeResourceAllocation

- Cree un clúster con Spark instalado y `maximizeResourceAllocation` establecido en true utilizando el AWS CLI archivo almacenado en Amazon S3`myConfig.json`, haciendo referencia a un archivo.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

`myConfig.json`:

```
[
  {
    "Classification": "spark",
    "Properties": {
```

```
      "maximizeResourceAllocation": "true"
    }
  }
]
```

> **Note**
>
> Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Migración de Apache Log4j 1.x a Log4j 2.x

Las versiones 3.2.x y anteriores de Apache Spark utilizan la versión heredada de Apache Log4j 1.x y el archivo `log4j.properties` para configurar Log4j en los procesos de Spark. Las versiones 3.3.0 y posteriores de Apache Spark utilizan Apache Log4j 2.x y el archivo `log4j2.properties` para configurar Log4j en los procesos de Spark.

Si ha configurado Apache Spark Log4j con una versión de Amazon EMR anterior a la 6.8.0, debe eliminar la clasificación de configuración `spark-log4j` heredada y completar la migración al formato de clave y clasificación de configuración `spark-log4j2` para poder actualizar a la versión 6.8.0 o posteriores de Amazon EMR. La clasificación `spark-log4j` heredada provoca un error `ValidationException` en la creación del clúster en las versiones 6.8.0 y posteriores de Amazon EMR. No se le cobrará nada en caso de un error relacionado con la incompatibilidad de Log4j, pero deberá eliminar la clasificación de configuración `spark-log4j` extinta para continuar.

Para obtener más información sobre la migración de Apache Log4j 1.x a Log4j 2.x, consulte Apache Log4j Migration Guide y Spark Log4j 2 Template en GitHub.

> **Note**
>
> Con Amazon EMR, Apache Spark usa un archivo `log4j2.properties` en lugar del archivo .xml descrito en Apache Log4j Migration Guide. Además, no recomendamos utilizar el método de puente de Log4j 1.x para llevar a cabo la conversión a Log4j 2.x.

Optimización del desempeño de Spark

Amazon EMR proporciona múltiples características de optimización del rendimiento para Spark. En este tema se explica cada característica de optimización en detalle.

Para obtener más información acerca de cómo definir la configuración de Spark, consulte Configurar Spark.

Ejecución de consultas adaptativas

La ejecución de consultas adaptativas es un marco para reoptimizar los planes de consultas en función de las estadísticas del tiempo de ejecución. A partir de Amazon EMR 5.30.0, las siguientes optimizaciones de ejecución de consultas adaptativas de Apache Spark 3 están disponibles en el tiempo de ejecución de Apache Amazon EMR para Spark 2.

- Conversión de unión adaptativa
- Fusión adaptativa de particiones aleatorias

Conversión de unión adaptativa

La conversión por combinación adaptativa mejora el rendimiento de las consultas al convertir sort-merge-join las operaciones en broadcast-hash-joins operaciones basadas en el tamaño del tiempo de ejecución de las etapas de la consulta. Broadcast-hash-joinstienden a funcionar mejor cuando un lado de la unión es lo suficientemente pequeño como para transmitir su salida de manera eficiente entre todos los ejecutores, lo que evita la necesidad de mezclar el intercambio y ordenar ambos lados de la combinación. La conversión adaptativa de uniones amplía la gama de casos en los que Spark funciona automáticamente. broadcast-hash-joins

Esta característica está habilitada de forma predeterminada. Para deshabilitarla, puede establecer `spark.sql.adaptive.enabled` en `false`, lo que también deshabilita el marco de ejecución de consultas adaptativas. Spark decide convertir a en sort-merge-join a broadcast-hash-join cuando la estadística del tamaño del tiempo de ejecución de uno de los lados de la unión no supera `spark.sql.autoBroadcastJoinThreshold` el valor predeterminado de 10.485.760 bytes (10 MiB).

Fusión adaptativa de particiones aleatorias

La fusión adaptativa de particiones aleatorias mejora el rendimiento de las consultas al fusionar pequeñas particiones aleatorias contiguas para evitar la sobrecarga que supone tener demasiadas

tareas pequeñas. Esto le permite configurar un mayor número de particiones aleatorias iniciales por adelantado, que luego se reduce en tiempo de ejecución hasta el tamaño deseado, lo que aumenta las posibilidades de tener particiones aleatorias distribuidas de manera más uniforme.

Esta característica está habilitada de forma predeterminada a menos que `spark.sql.shuffle.partitions` se establezca de forma explícita. Para habilitarla, se puede establecer `spark.sql.adaptive.coalescePartitions.enabled` en `true`. Tanto el número inicial de particiones aleatorias como el tamaño de las particiones de destino se pueden ajustar con las propiedades `spark.sql.adaptive.coalescePartitions.minPartitionNum` y `spark.sql.adaptive.advisoryPartitionSizeInBytes`, respectivamente. Consulta la siguiente tabla para obtener más información sobre las propiedades de Spark relacionadas con esta característica.

Propiedades de las particiones de la fusión adaptativa de Spark

Propiedad	Valor predeterminado	Descripción
`spark.sql.adaptive.coalescePartitions.enabled`	true, a menos que `spark.sql.shuffle.partitions` se establezca de forma explícita	Si el valor es true y spark.sql.adaptive.enabled es true, Spark fusiona las particiones aleatorias contiguas según el tamaño de destino (especificado mediante `spark.sql.adaptive.advisoryPartitionSizeInBytes`) para evitar demasiadas tareas pequeñas.
`spark.sql.adaptive.advisoryPartitionSizeInBytes`	64 MB	Tamaño recomendado en bytes de la partición aleatoria al fusionarse. Esta configuración solo tiene efecto cuando `spark.sql.adaptive.enabled` y `spark.sql.adaptive.coalescePartitions.enabled` son `true`.

Propiedad	Valor predeterminado	Descripción
`spark.sql.adaptive.coalescePartitions.minPartitionNum`	25	Número mínimo de particiones aleatorias después de la fusión. Esta configuración solo tiene efecto cuando `spark.sql.adaptive.enabled` y `spark.sql.adaptive.coalescePartitions.enabled` son `true`.
`spark.sql.adaptive.coalescePartitions.initialPartitionNum`	1 000	Número inicial de particiones aleatorias antes de la fusión. Esta configuración solo tiene efecto cuando `spark.sql.adaptive.enabled` y `spark.sql.adaptive.coalescePartitions.enabled` son `true`.

Eliminación dinámica de particiones

La reducción dinámica de particiones mejora el rendimiento seleccionando con mayor precisión particiones específicas de una tabla que deben leerse y procesarse para una consulta específica. Al reducir la cantidad de datos leídos y procesados, se ahorra bastante tiempo en la ejecución de trabajos. Con Amazon EMR 5.26.0, esta característica está habilitada de forma predeterminada. Con Amazon EMR 5.24.0 y 5.25.0, puede establecer la propiedad `spark.sql.dynamicPartitionPruning.enabled` de Spark desde Spark o al crear clústeres para habilitar esta característica.

Propiedades de las particiones de la eliminación dinámica de particiones de Spark

Propiedad	Valor predeterminado	Descripción
`spark.sql.dynamicPartitionPruning.enabled`	`true`	Si es true, habilite la eliminación dinámica de particiones.
`spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`	`true`	Si es `true`, Spark realiza una comprobación preventiva antes de ejecutar las consultas para garantizar que la reutilización de los intercambios de transmisiones en filtros de eliminación dinámica no se vea afectada por reglas de preparación posteriores (como las reglas de columnas definidas por el usuario). Cuando la reutilización se ve afectada y esta configuración es `true`, Spark elimina los filtros de eliminación dinámica afectados para evitar problemas de rendimiento y corrección. Pueden surgir problemas de corrección cuando el intercambio de transmisión del filtro de eliminación dinámica ofrece resultados diferentes e incoherentes a partir del intercambio de transmisión de la operación de unión correspondiente. Definir esta configuración en `false` debe hacerse con

Propiedad	Valor predeterminado	Descripción
		precaución, ya que permite evitar situaciones como, por ejemplo, cuando la reutilización se ve afectada por reglas de columnas definidas por el usuario. Cuando la ejecución de consultas adaptativas está habilitada, siempre se aplica la reutilización de transmisiones.

Esta optimización mejora las capacidades existentes de Spark 2.4.2, que solo admiten la postergación de predicados estáticos que se pueden resolver durante la planificación.

A continuación, se muestran ejemplos de postergación de predicados estáticos en Spark 2.4.2.

```
partition_col = 5

partition_col IN (1,3,5)

partition_col between 1 and 3

partition_col = 1 + 3
```

La reducción dinámica de particiones permite que el motor de Spark deduzca de forma dinámica durante el tiempo de ejecución qué particiones deben leerse y cuáles pueden eliminarse de forma segura. Por ejemplo, la siguiente consulta implica dos tablas: la tabla `store_sales` que contiene todas las ventas totales para todas las tiendas y se particiona por región y la tabla `store_regions` que contiene un mapeo de las regiones de cada país. Las tablas contienen datos sobre tiendas distribuidas por todo el mundo, pero solo realizan consultas en los datos de América del Norte.

```
select ss.quarter, ss.region, ss.store, ss.total_sales
from store_sales ss, store_regions sr
where ss.region = sr.region and sr.country = 'North America'
```

Sin la reducción de particiones dinámicas, esta consulta leerá todas las regiones antes de filtrar el subconjunto de regiones que coinciden con los resultados de la subconsulta. Con la reducción

dinámica de particiones, esta consulta leerá y procesará solo las particiones de las regiones que se devuelven en la subconsulta. Esto ahorra tiempo y recursos ya que lee menos datos del almacenamiento y procesa menos registros.

Aplanamiento de subconsultas escalares

Esta optimización mejora el rendimiento de las consultas que tienen subconsultas escalares a través de la misma tabla. Con Amazon EMR 5.26.0, esta característica está habilitada de forma predeterminada. Con Amazon EMR 5.24.0 y 5.25.0, puede establecer la propiedad `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled` de Spark desde Spark o al crear clústeres para habilitarlo. Cuando esta propiedad se establece en true, el optimizador de consultas aplana subconsultas escalares de agregación que utilizan la misma relación en caso posible. Las consultas escalares se aplanan empujando cualquier predicado presente en la subconsulta en las funciones de agregación y, a continuación, realiza una agregación, con todas las funciones de agregación, por relación.

A continuación se presenta un ejemplo de una consulta que se beneficiará de esta optimización.

```
select (select avg(age) from students                    /* Subquery 1 */
                where age between 5 and 10) as group1,
       (select avg(age) from students                    /* Subquery 2 */
                where age between 10 and 15) as group2,
       (select avg(age) from students                    /* Subquery 3 */
                where age between 15 and 20) as group3
```

La optimización vuelve a escribir la consulta anterior como:

```
select c1 as group1, c2 as group2, c3 as group3
from (select avg (if(age between 5 and 10, age, null)) as c1,
             avg (if(age between 10 and 15, age, null)) as c2,
             avg (if(age between 15 and 20, age, null)) as c3 from students);
```

Tenga en cuenta que la consulta reescrita lee la tabla solo una vez, y los predicados de las tres subconsultas se envían a la función `avg`.

DISTINCT antes de INTERSECT

Esta optimización optimiza uniones cuando se utiliza INTERSECT. Con Amazon EMR 5.26.0, esta característica está habilitada de forma predeterminada. Con Amazon EMR 5.24.0 y 5.25.0, puede establecer la propiedad `spark.sql.optimizer.distinctBeforeIntersect.enabled` de

Spark desde Spark o al crear clústeres para habilitarlo. Las consultas con INTERSECT se convierten automáticamente para utilizar una semiunión pendiente. Cuando esta propiedad se establece en true, el optimizador de consultas coloca el operador DISTINCT en los elementos secundarios de INTERSECT si detecta que el operador DISTINCT puede hacer que el polo izquierdo se una a en lugar de a. BroadcastHashJoin SortMergeJoin

A continuación se presenta un ejemplo de una consulta que se beneficiará de esta optimización.

```
(select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
intersect
(select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
```

Sin habilitar esta propiedad `spark.sql.optimizer.distinctBeforeIntersect.enabled`, la consulta se reescribirá tal y como se indica a continuación.

```
select distinct brand from
  (select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
left semi join
   (select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
 on brand <=> cs_brand
```

Cuando se habilita esta propiedad `spark.sql.optimizer.distinctBeforeIntersect.enabled`, la consulta se reescribirá tal y como se indica a continuación.

```
select brand from
  (select distinct item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
left semi join
   (select distinct item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
 on brand <=> cs_brand
```

Unión de filtros de Bloom

Esta optimización puede mejorar el rendimiento de algunas uniones efectuando un filtrado previo de un lado de una unión mediante un filtro de Bloom generado a partir de los valores

del otro lado de la unión. Con Amazon EMR 5.26.0, esta característica está habilitada de forma predeterminada. Con Amazon EMR 5.25.0, puede establecer la propiedad de Spark `spark.sql.bloomFilterJoin.enabled` en `true` desde Spark o al crear clústeres para habilitar esta característica.

A continuación se muestra un ejemplo de consulta que puede beneficiarse de un filtro de Bloom.

```
select count(*)
from sales, item
where sales.item_id = item.id
and item.category in (1, 10, 16)
```

Cuando esta característica está habilitada, el filtro de Bloom se crea a partir de todos los identificadores de elemento cuya categoría se encuentra en el conjunto de categorías que se están consultando. Durante el análisis de la tabla de ventas, el filtro de Bloom se utiliza para determinar qué ventas corresponden a artículos que definitivamente no están en el conjunto definido por el filtro de Bloom. De este modo, estas ventas identificadas pueden excluirse lo antes posible.

Reordenación de unión optimizada

Esta optimización puede mejorar el rendimiento de las consultas al reordenar las uniones en las que intervienen tablas con filtros. Con Amazon EMR 5.26.0, esta característica está habilitada de forma predeterminada. Con Amazon EMR 5.25.0, puede establecer el parámetro de configuración `spark.sql.optimizer.sizeBasedJoinReorder.enabled` de Spark en true para habilitar esta característica. El comportamiento predeterminado en Spark es unir las tablas de izquierda a derecha, como se indica en la consulta. Esta estrategia puede perder oportunidades para ejecutar primero uniones más pequeñas con filtros, con el fin de beneficiarse de uniones más costosas más tarde.

La consulta de ejemplo que aparece a continuación informa de todos los artículos devueltos de todas las tiendas de un país. Sin la reordenación de unión optimizada, Spark une primero las dos tablas grandes `store_sales` y `store_returns`; a continuación, las une con `store` y, finalmente, con `item`.

```
select ss.item_value, sr.return_date, s.name, i.desc,
from store_sales ss, store_returns sr, store s, item i
where ss.id = sr.id and ss.store_id = s.id and ss.item_id = i.id
and s.country = 'USA'
```

Con la reordenación de unión optimizada, Spark primero une `store_sales` con `store`, ya que `store` tiene un filtro y es más pequeño que `store_returns` y `broadcastable`. Luego Spark se une con `store_returns` y, finalmente, con `item`. Si `item` tuviera un filtro y se pudiera emitir, también se podría usar en la reordenación, lo que provocaría que `store_sales` se uniera con `store`, después con `item` y, finalmente, con `store_returns`.

Almacenamiento en caché de fragmentos de resultados de Spark

La versión 6.6.0 y posteriores de Amazon EMR incluyen la característica opcional Almacenamiento en caché de fragmentos de resultados de Spark, que almacena en caché automáticamente los fragmentos de resultados. Estos fragmentos de resultados forman parte de los resultados de subárboles de consultas que se almacenan en el bucket de Amazon S3 que elija. Los fragmentos de resultados de las consultas almacenados se reutilizan en las siguientes ejecuciones de consultas, lo que se traduce en consultas más rápidas.

Result Fragment Caching analiza tus consultas de Spark y almacena en caché los fragmentos de resultados aptos en la ubicación S3 que especifiques. En las siguientes ejecuciones de consultas, los fragmentos de resultados de consultas que se pueden utilizar se detectan automáticamente y se recuperan de S3. El Almacenamiento en caché de fragmentos de resultados es diferente del almacenamiento en caché de conjuntos de resultados, en el que las consultas posteriores tienen que coincidir exactamente con la consulta original para obtener resultados de la caché. Cuando se utiliza para consultas que se dirigen repetidamente a un subconjunto estático de los datos, el Almacenamiento en caché de fragmentos de resultados acelera considerablemente el rendimiento.

Considere la siguiente consulta, que cuenta los pedidos hasta el año 2022:

```
select
    l_returnflag,
    l_linestatus,
    count(*) as count_order
from
    lineitem
where
    l_shipdate <= current_date
    and year(l_shipdate) == '2022'
group by
    l_returnflag,
    l_linestatus
```

A medida que pasa el tiempo, esta consulta debe ejecutarse todos los días para informar sobre las ventas totales del año. Sin el Almacenamiento en caché de fragmentos de resultados, los resultados de todos los días del año deberán volver a calcularse todos los días. La consulta se volverá más lenta con el tiempo y será más lenta al final del año, cuando sea necesario volver a calcular los 365 días de resultados.

Al activar el Almacenamiento en caché de fragmentos de resultados, se utilizan los resultados de todos los días anteriores del año guardados en la caché. Cada día, la característica debe volver a calcular los resultados de un solo día. Una vez que la característica calcula el fragmento de resultados, lo almacena en caché. Como resultado, los tiempos de consultas con la caché habilitada son rápidos y constantes para cada consulta posterior.

Habilitación del Almacenamiento en caché de fragmentos de resultados de Spark

Para habilitar el Almacenamiento en caché de fragmentos de resultados de Spark, lleve a cabo los siguientes pasos:

1. Cree un depósito de caché en Amazon S3 y autorice el read/write acceso a EMRFS. Para obtener más información, consulte Autorización de acceso a los datos de EMRFS en Amazon S3.
2. Establezca la configuración de Spark de Amazon EMR para habilitar la característica.

```
spark.subResultCache.enabled = true
spark.subResultCache.fs.root.path = s3://&example-s3-bucket;/cache_dir/
```

3. Habilite la administración del ciclo de vida de S3 para que el bucket limpie automáticamente los archivos de caché.
4. Si lo desea, configure las maxBufferSize propiedades reductionRationThreshold y para ajustar aún más la función.

```
spark.sql.subResultCache.reductionRatioThreshold
spark.sql.subResultCache.maxBufferSize
```

Consideraciones al utilizar el Almacenamiento en caché de fragmentos de resultados

El ahorro de costos al utilizar los resultados ya almacenados en caché en Amazon S3 en lugar de volver a calcularlos aumenta con el número de veces que se pueden utilizar los mismos resultados almacenados en caché. Las consultas con análisis de tablas de gran tamaño seguidas de filtros o agregaciones de hash que reduzcan el tamaño de los resultados en un factor mínimo de 8 (es decir, una proporción de al menos 8:1 entre tamaño de entrada y resultados) son las que más se beneficiarán de esta característica. Cuanto mayor sea la proporción de reducción entre la entrada y los resultados, mayor será la relación entre costos y beneficios. Las consultas con proporciones de reducción más bajas, pero que contengan pasos costosos de computación entre el análisis de la tabla y el filtro o las agregaciones, también se beneficiarán, siempre que el costo de producir los resultados sea mayor que el costo de obtenerlos de Amazon S3. De forma predeterminada, el Almacenamiento en caché de fragmentos de resultados solo se aplica cuando detecta que la proporción de reducción será de al menos 8:1.

Cuando las consultas reutilizan repetidamente los resultados almacenados en caché, las ventajas de esta característica son mayores. Las consultas en intervalos continuos e incrementales son buenos ejemplos. Por ejemplo, una consulta continua de 30 días que ya se haya ejecutado durante 29 días solo necesitaría extraer una trigésima parte de los datos de destino de su origen de entrada original y utilizaría fragmentos de resultados almacenados en caché de los 29 días anteriores. Una consulta de intervalo incremental se beneficia aún más, ya que el inicio del intervalo permanece fijo: cada vez que se invoca la consulta, un porcentaje menor del procesamiento requerirá la lectura del origen de entrada.

A continuación se indican algunas consideraciones adicionales a la hora de utilizar el Almacenamiento en caché de fragmentos de resultados:

- Las consultas que no se dirijan a los mismos datos con los mismos fragmentos de consulta tendrán una tasa de aciertos de caché baja, por lo que no se beneficiarán de esta característica.
- Las consultas con proporciones de reducción bajas que no contengan pasos de computación costosos darán como resultado resultados almacenados en caché que son aproximadamente tan caros de leer como de procesar inicialmente.
- La primera consulta siempre mostrará una regresión menor debido al costo de escribir en la caché.
- La característica Almacenamiento en caché de fragmentos de resultados funciona exclusivamente con archivos Parquet. No se admite ningún otro formato de archivo.

- Los búferes de la característica Almacenamiento en caché de fragmentos de resultados solo intentarán almacenar en caché los análisis con tamaños de división de archivos de 128 MB o más. Con la configuración predeterminada de Spark, el Almacenamiento en caché de fragmentos de resultados se deshabilitará si el tamaño del análisis (el tamaño total de todos los archivos que se analizan) dividido por el número de núcleos ejecutores es inferior a 128 MB. Cuando se establezca alguna de las configuraciones de Spark que se enumeran a continuación, el tamaño de división de archivos será:

```
min(maxPartitionBytes, max(openCostInBytes, scan size / minPartitionNum))
```

  - spark.sql. leafNodeDefaultParalelismo (el valor predeterminado es spark.default.parallelism)
  - spark.sql.files. minPartitionNum (el valor predeterminado es spark.sql. leafNodeDefaultParalelismo)
  - spark.sql.files. openCostInBytes
  - spark.sql.files. maxPartitionBytes
- La característica Almacenamiento en caché de fragmentos de resultados almacena en caché según la granularidad de la partición de RDD. La proporción de reducción descrita anteriormente, que de forma predeterminada es 8:1, se evalúa por partición de RDD. Las cargas de trabajo con proporciones de reducción por RDD superiores e inferiores a 8:1 pueden obtener menores ventajas de rendimiento que las cargas de trabajo con proporciones de reducción por RDD que son sistemáticamente inferiores a 8:1.
- La característica Almacenamiento en caché de fragmentos de resultados utiliza un búfer de escritura de 16 MB de forma predeterminada para cada partición de RDD que se almacena en caché. Si se van a almacenar en caché más de 16 MB por partición de RDD, el costo de determinar que no es posible realizar una escritura puede provocar una regresión del rendimiento.
- Si bien de forma predeterminada el Almacenamiento en caché de fragmentos de resultados no intentará almacenar en caché los resultados de las particiones de RDD con una proporción de reducción inferior a 8:1 y limitará su búfer de escritura a 16 MB, ambos valores se pueden ajustar mediante las siguientes configuraciones:

```
spark.sql.subResultCache.reductionRatioThreshold (default: 8.0)
spark.sql.subResultCache.maxBufferSize (default: 16MB, max: 64MB)
```

- Varios clústeres que utilizan la misma versión de Amazon EMR pueden compartir la misma ubicación de caché. Para garantizar la exactitud de los resultados, el Almacenamiento en caché de

fragmentos de resultados no utilizará los resultados almacenados en caché escritos por distintas versiones de Amazon EMR.

- El almacenamiento en caché de fragmentos de resultados se desactivará automáticamente en los casos de uso de Spark Streaming o cuando se RecordServer utilice Apache Ranger o AWS Lake Formation se utilice.
- La caché de fragmentos de resultados read/writes utiliza buckets EMRFS/ S3A y Amazon S3. Se admite el cifrado CSE (solo con EMRFS), SSE S3/ SSE KMS. A modo de contexto, S3A proporciona una implementación de Hadoop para permitir que un clúster lea y escriba datos desde y hacia Amazon S3. Tenga en cuenta que la compatibilidad con S3A está disponible con EMR-7.4.0 y versiones posteriores.

Uso de Nvidia RAPIDS Accelerator de Nvidia para Spark

Con Amazon EMR versión 6.2.0 y versiones posteriores, puede utilizar el complemento RAPIDS Accelerator para Apache Spark de Nvidia para acelerar Spark mediante tipos de instancias de unidades de procesamiento EC2 gráfico (GPU). RAPIDS Accelerator acelerará mediante GPU sus canalizaciones de ciencia de datos de Apache Spark 3.0 sin cambios de código y acelerará el procesamiento de datos y el entrenamiento de modelos, a la vez que reducirá considerablemente los costos de infraestructura.

En las siguientes secciones lo guiaremos a través de la configuración de su clúster de EMR para usar el complemento Spark-RAPIDS para Spark.

Elección de tipos de instancias

Para usar el complemento Spark-rapids de Nvidia para Spark, los grupos de instancias principales y de tareas deben usar tipos de instancias de EC2 GPU que cumplan con los requisitos de hardware de Spark-rapids. Para ver una lista completa de los tipos de instancias de GPU compatibles con Amazon EMR, consulte Tipos de instancias admitidos en la Guía de administración de Amazon EMR. El tipo de instancia del grupo de instancias principal puede ser de tipo GPU o no, pero no se admiten los tipos de instancias de ARM.

Definición de las configuraciones de aplicaciones para un clúster

1. Permitir que Amazon EMR instale los complementos en su nuevo clúster

Para instalar los complementos, proporcione la siguiente configuración al crear el clúster:

```
{
 "Classification":"spark",
 "Properties":{
  "enableSparkRapids":"true"
 }
}
```

2. Configurar YARN para usar GPU

Para obtener más información sobre cómo usar GPU en YARN, consulte Using GPU on YARN en la documentación de Apache Hadoop. En los siguientes ejemplos, se muestran ejemplos de configuraciones de YARN para las versiones 6.x y 7.x de Amazon EMR:

Amazon EMR 7.x

Ejemplo de configuración de YARN para Amazon EMR 7.x

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
```

```
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/spark-rapids-cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
}
```

Amazon EMR 6.x

Ejemplo de configuración de YARN para Amazon EMR 6.x

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
```

```
            "Properties":{
                "root":"/sys/fs/cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
}
```

3. Configurar Spark para usar RAPIDS

Estas son las configuraciones necesarias para que Spark pueda usar el complemento RAPIDS:

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.plugins":"com.nvidia.spark.SQLPlugin",
  "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/gpu/
getGpusResources.sh",
  "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-linux/lib:/usr/
local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/cuda/lib:/usr/
local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/native:/docker/usr/
lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native"
 }
}
```

XGBoost4La [biblioteca J-Spark](#) de la XGBoost documentación también está disponible cuando el complemento Spark RAPIDS está activado en el clúster. Puedes usar la siguiente configuración para integrarla XGBoost con tu trabajo de Spark:

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-spark_3.0-1.4.2-0.3.0.jar"
 }
}
```

Para obtener más configuraciones de Spark que pueda usar para ajustar un clúster de EMR acelerado mediante GPU, consulte [Rapids Accelerator for Apache Spark tuning guide](#) en la documentación de Nvidia.github.io.

4. Configurar el programador de capacidad de YARN

`DominantResourceCalculator` debe configurarse para habilitar la programación y el aislamiento de GPU. Para obtener más información, consulte Using GPU on YARN en la documentación de Apache Hadoop.

```
{
 "Classification":"capacity-scheduler",
 "Properties":{
  "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
 }
}
```

5. Crear un archivo JSON para incluir todas las configuraciones

Puede crear un archivo JSON que contenga la configuración para usar el complemento RAPIDS para su clúster de Spark. Proporcione el archivo más adelante, al lanzar el clúster.

Puede almacenar el archivo de forma local o en &S3. Para obtener más información sobre cómo proporcionar configuraciones de aplicaciones para sus clústeres, consulte Configuración de aplicaciones.

Utilice los siguientes archivos de ejemplo como plantillas para crear sus propias configuraciones.

Amazon EMR 7.x

Archivo **`my-configurations.json`** de ejemplo para Amazon EMR 7.x

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
```

```
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
    {
        "Classification":"container-executor",
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/spark-rapids-cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
            "spark.task.cpus":"1",
```

```
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
    },
    {
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
]
```

Amazon EMR 6.x

Archivo **`my-configurations.json`** de ejemplo para Amazon EMR 6.x

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
```

```
    {
        "Classification":"container-executor",
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/sys/fs/cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
            "spark.task.cpus":"1",
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
```

```
    },
    {
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
]
```

Adición de una acción de arranque para el clúster

Para obtener más información sobre cómo proporcionar scripts de acciones de arranque al crear el clúster, consulte Conceptos básicos de las acciones de arranque en la Guía de administración de Amazon EMR.

En los siguientes scripts de ejemplo, se muestra cómo crear un archivo de acciones de arranque para Amazon EMR 6.x y 7.x:

Amazon EMR 7.x

Archivo **my-bootstrap-action.sh** de ejemplo para Amazon EMR 7.x

Para utilizar YARN con el fin de administrar los recursos de GPU con las versiones 7.x de Amazon EMR, debe realizar el montaje de forma manual de CGroup v1 en el clúster. Puede hacerlo con un script de acción de arranque, como se muestra en este ejemplo.

```
#!/bin/bash
set -ex

sudo mkdir -p /spark-rapids-cgroup/devices
sudo mount -t cgroup -o devices cgroupv1-devices /spark-rapids-cgroup/devices
sudo chmod a+rwx -R /spark-rapids-cgroup
```

Amazon EMR 6.x

Archivo **my-bootstrap-action.sh** de ejemplo para Amazon EMR 6.x

Para las versiones 6.x de Amazon EMR, debe abrir los permisos de CGroup para YARN en el clúster. Puede hacerlo con un script de acciones de arranque, como se muestra en este ejemplo.

```
#!/bin/bash
set -ex

sudo chmod a+rwx -R /sys/fs/cgroup/cpu,cpuacct
sudo chmod a+rwx -R /sys/fs/cgroup/devices
```

Lanzamiento del clúster

El último paso consiste en lanzar el clúster con las configuraciones de clúster mencionadas anteriormente. A continuación, se muestra un comando de ejemplo para lanzar un clúster desde la CLI de Amazon EMR:

```
 aws emr create-cluster \
--release-label emr-7.9.0 \
--applications Name=Hadoop Name=Spark \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=my-key-pair,InstanceProfile=EMR_EC2_DefaultRole \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m4.4xlarge \

                  InstanceGroupType=CORE,InstanceCount=1,InstanceType=g4dn.2xlarge \

                  InstanceGroupType=TASK,InstanceCount=1,InstanceType=g4dn.2xlarge \
--configurations file:///my-configurations.json \
--bootstrap-actions Name='My Spark Rapids Bootstrap action',Path=s3://amzn-s3-demo-
bucket/my-bootstrap-action.sh
```

Acceso al intérprete de comandos de Spark

El intérprete de comandos de Spark se basa en el Scala REPL (Read-Eval-Print-Loop). Le permite crear programas de Spark de forma interactiva y enviar trabajos al marco. Para acceder al intérprete de comandos de Spark, puede conectarse al nodo principal con SSH e invocar `spark-shell`. Para obtener más información sobre cómo conectarse al nodo principal, consulte Conectarse al nodo principal mediante SSH en la Guía de administración de Amazon EMR. En los siguientes ejemplos se utilizan registros de acceso de Apache HTTP Server almacenados en Amazon S3.

> **Note**
>
> El bucket de estos ejemplos está disponible para los clientes que pueden acceder a la región Este de EE. UU. (Norte de Virginia).

Por defecto, el shell Spark crea su propio SparkContextobjeto llamado`sc`. Puedes usar este contexto si es necesario en el REPL. sqlContexttambién está disponible en la carcasa y es un HiveContext.

Example Uso del intérprete de comandos de Spark para contar cuántas veces aparece una cadena en un archivo almacenado en Amazon S3

En este ejemplo, se utiliza `sc` para leer un archivo de texto almacenado en Amazon S3.

```
scala> sc
res0: org.apache.spark.SparkContext = org.apache.spark.SparkContext@404721db

scala> val textFile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/
impressions/dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark crea el textFile y la estructura de datos asociada. A continuación, el ejemplo determina la cantidad de líneas en el archivo de registro con la cadena "cartoonnetwork.com":

```
scala> val linesWithCartoonNetwork = textFile.filter(line =>
 line.contains("cartoonnetwork.com")).count()
linesWithCartoonNetwork: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[2] at
 filter at <console>:23
<snip>
<Spark program runs>
scala> linesWithCartoonNetwork
res2: Long = 9
```

Example Uso del intérprete de comandos de Spark basado en Python para contar cuántas veces aparece una cadena en un archivo almacenado en Amazon S3

Spark también incluye un shell basado en Python, `pyspark`, que puede utilizar para crear prototipos de programas de Spark escritos en Python. Al igual que con`spark-shell`, invoke `pyspark` en el nodo principal; también tiene el mismo SparkContextobjeto.

```
>>> sc
```

```
<pyspark.context.SparkContext object at 0x7fe7e659fa50>
>>> textfile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/impressions/
dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark crea el textFile y la estructura de datos asociada. A continuación, el ejemplo determina la cantidad de líneas en el archivo de registro con la cadena "cartoonnetwork.com".

```
>>> linesWithCartoonNetwork = textfile.filter(lambda line: "cartoonnetwork.com" in
 line).count()
15/06/04 17:12:22 INFO lzo.GPLNativeCodeLoader: Loaded native gpl library from the
 embedded binaries
15/06/04 17:12:22 INFO lzo.LzoCodec: Successfully loaded & initialized native-lzo
 library [hadoop-lzo rev EXAMPLE]
15/06/04 17:12:23 INFO fs.EmrFileSystem: Consistency disabled, using
 com.amazon.ws.emr.hadoop.fs.s3n.S3NativeFileSystem as filesystem implementation
<snip>
<Spark program continues>
>>> linesWithCartoonNetwork
9
```

Utilice Amazon SageMaker Spark para el aprendizaje automático

Cuando se utiliza la versión 5.11.0 y posteriores de Amazon EMR, el componente `aws-sagemaker-spark-sdk` se instala junto con Spark. Este componente instala Amazon SageMaker Spark y las dependencias asociadas para la integración de Spark con Amazon. SageMaker Puedes usar Amazon SageMaker Spark para crear canalizaciones de aprendizaje automático (ML) de Spark mediante Amazon SageMaker Stages. Para obtener más información, consulta el README de Amazon SageMaker Spark sobre el GitHub uso de Apache Spark con Amazon SageMaker en la Guía para SageMaker desarrolladores de Amazon.

Escritura de una aplicación de Spark

Las aplicaciones de Spark se pueden escribir en Scala, Java o Python. Hay varios ejemplos de aplicaciones de Spark en el tema Spark examples en la documentación de Apache Spark. El ejemplo Estimación de Pi que se muestra a continuación en las tres aplicaciones compatibles de forma nativa. También puede ver ejemplos completos en `$SPARK_HOME/examples` y en GitHub. Para obtener más información sobre cómo compilar JARs para Spark, consulta el tema de inicio rápido en la documentación de Apache Spark.

Scala

Para evitar problemas de compatibilidad con Scala, le sugerimos que utilice las dependencias de Spark para la versión correcta de Scala cuando compile una aplicación de Spark para un clúster de Amazon EMR. La versión de Scala que debe usar depende de la versión de Spark instalada en el clúster. Por ejemplo, la versión 5.30.1 de Amazon EMR usa Spark 2.4.5, que se basa en Scala 2.11. Si su clúster utiliza la versión 5.30.1 de Amazon EMR, utilice las dependencias de Spark para Scala 2.11. Para obtener más información sobre las versiones de Scala que utiliza Spark, consulte la documentación de Apache Spark.

```
package org.apache.spark.examples
import scala.math.random
import org.apache.spark._

/** Computes an approximation to pi */
object SparkPi {
  def main(args: Array[String]) {
    val conf = new SparkConf().setAppName("Spark Pi")
    val spark = new SparkContext(conf)
    val slices = if (args.length > 0) args(0).toInt else 2
    val n = math.min(100000L * slices, Int.MaxValue).toInt // avoid overflow
    val count = spark.parallelize(1 until n, slices).map { i =>
      val x = random * 2 - 1
      val y = random * 2 - 1
      if (x*x + y*y < 1) 1 else 0
    }.reduce(_ + _)
    println("Pi is roughly " + 4.0 * count / n)
    spark.stop()
  }
}
```

Java

```
package org.apache.spark.examples;

import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaRDD;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.api.java.function.Function;
import org.apache.spark.api.java.function.Function2;
```

```
import java.util.ArrayList;
import java.util.List;

/**
 * Computes an approximation to pi
 * Usage: JavaSparkPi [slices]
 */
public final class JavaSparkPi {

  public static void main(String[] args) throws Exception {
    SparkConf sparkConf = new SparkConf().setAppName("JavaSparkPi");
    JavaSparkContext jsc = new JavaSparkContext(sparkConf);

    int slices = (args.length == 1) ? Integer.parseInt(args[0]) : 2;
    int n = 100000 * slices;
    List<Integer> l = new ArrayList<Integer>(n);
    for (int i = 0; i < n; i++) {
      l.add(i);
    }

    JavaRDD<Integer> dataSet = jsc.parallelize(l, slices);

    int count = dataSet.map(new Function<Integer, Integer>() {
      @Override
      public Integer call(Integer integer) {
        double x = Math.random() * 2 - 1;
        double y = Math.random() * 2 - 1;
        return (x * x + y * y < 1) ? 1 : 0;
      }
    }).reduce(new Function2<Integer, Integer, Integer>() {
      @Override
      public Integer call(Integer integer, Integer integer2) {
        return integer + integer2;
      }
    });

    System.out.println("Pi is roughly " + 4.0 * count / n);

    jsc.stop();
  }
}
```

Python

```
import argparse
import logging
from operator import add
from random import random

from pyspark.sql import SparkSession

logger = logging.getLogger(__name__)
logging.basicConfig(level=logging.INFO, format="%(levelname)s: %(message)s")


def calculate_pi(partitions, output_uri):
    """
    Calculates pi by testing a large number of random numbers against a unit circle
    inscribed inside a square. The trials are partitioned so they can be run in
    parallel on cluster instances.

    :param partitions: The number of partitions to use for the calculation.
    :param output_uri: The URI where the output is written, typically an Amazon S3
                       bucket, such as 's3://example-bucket/pi-calc'.
    """

    def calculate_hit(_):
        x = random() * 2 - 1
        y = random() * 2 - 1
        return 1 if x**2 + y**2 < 1 else 0

    tries = 100000 * partitions
    logger.info(
        "Calculating pi with a total of %s tries in %s partitions.", tries, partitions
    )
    with SparkSession.builder.appName("My PyPi").getOrCreate() as spark:
        hits = (
            spark.sparkContext.parallelize(range(tries), partitions)
            .map(calculate_hit)
            .reduce(add)
        )
        pi = 4.0 * hits / tries
        logger.info("%s tries and %s hits gives pi estimate of %s.", tries, hits, pi)
        if output_uri is not None:
```

```
            df = spark.createDataFrame([(tries, hits, pi)], ["tries", "hits", "pi"])
            df.write.mode("overwrite").json(output_uri)


if __name__ == "__main__":
    parser = argparse.ArgumentParser()
    parser.add_argument(
        "--partitions",
        default=2,
        type=int,
        help="The number of parallel partitions to use when calculating pi.",
    )
    parser.add_argument(
        "--output_uri", help="The URI where output is saved, typically an S3 bucket."
    )
    args = parser.parse_args()

    calculate_pi(args.partitions, args.output_uri)
```

Mejora del rendimiento de Spark con Amazon S3

Amazon EMR ofrece características para ayudar a optimizar el rendimiento cuando se utiliza Spark para consultar, leer y escribir datos guardados en Amazon S3.

S3 Select puede mejorar el rendimiento de las consultas para archivos JSON y CSV en algunas aplicaciones al “delegar” el procesamiento en Amazon S3.

El compilador optimizado para EMRFS para S3 es una alternativa a la OutputCommitterclase, que utiliza la función de cargas multiparte de EMRFS para mejorar el rendimiento al escribir archivos Parquet en Amazon S3 mediante Spark y Datasets. DataFrames

Temas

- Uso de S3 Select con Spark para mejorar el rendimiento de las consultas
- Uso del confirmador optimizado para S3 de EMRFS
- Uso del protocolo de confirmación optimizado para S3 de EMRFS
- Reintento de solicitudes de Amazon S3 con EMRFS

Uso de S3 Select con Spark para mejorar el rendimiento de las consultas

> **Important**
>
> Amazon S3 Select ya no está disponible para los nuevos clientes. Los clientes actuales de Amazon S3 Select pueden seguir utilizando la característica de la forma habitual. Más información

Con la versión 5.17.0 y posteriores de Amazon EMR, puede usar S3 Select con Spark en Amazon EMR. S3 Select permite que las aplicaciones recuperen únicamente un subconjunto de datos de un objeto. Para Amazon EMR, el clúster "delega" en Amazon S3 el trabajo de computación de filtrar grandes conjuntos de datos para su procesamiento, lo que puede mejorar el rendimiento en algunas aplicaciones y reduce la cantidad de datos transferidos entre Amazon EMR y Amazon S3.

S3 Select es compatible con los archivos JSON y CSV si se utilizan los valores `s3selectCSV` y `s3selectJSON` para especificar el formato de los datos. Para obtener más información y ejemplos, consulta Especificación de S3 Select en el código.

¿S3 Select es adecuado para mi aplicación?

Le recomendamos que realice un análisis comparativo sus aplicaciones con y sin S3 Select para ver si su uso puede ser adecuado para su aplicación.

Utilice las siguientes directrices para determinar si la aplicación es candidata a utilizar S3 Select:

- La consulta filtra más de la mitad del conjunto de datos original.
- La conexión de red entre Amazon S3 y el clúster de Amazon EMR tiene una buena velocidad de transferencia y ancho de banda disponible. Amazon S3 no comprime las respuestas HTTP, por lo que es probable que el tamaño de la respuesta aumente en el caso de los archivos de entrada comprimidos.

Consideraciones y limitaciones

- No se admiten el cifrado del servidor de Amazon S3 con claves de cifrado proporcionadas por el cliente (SSE-C) ni el cifrado del cliente.
- No se admite la propiedad `AllowQuotedRecordDelimiters`. Si se especifica esta propiedad, se producirá un error en la consulta.

- Solo se admiten archivos CSV y JSON en formato UTF-8. No se admiten las líneas múltiples. CSVs
- Solo se admiten los archivos sin comprimir o gzip.
- No se admiten opciones de JSON y CSV en Spark como, por ejemplo `nanValue`, `positiveInf`, `negativeInf` y las opciones relacionadas con los registros dañados (por ejemplo, failfast y el modo dropmalformed).
- El uso de comas (,) dentro de los decimales no se admite. Por ejemplo, `10,000` no se admite, pero sí se admite `10000`.
- No se admiten los caracteres de comentario en la última línea.
- Las líneas vacías al final de un archivo no se procesan.
- Los siguientes filtros no se delegan en Amazon S3:
  - Funciones de agregación como `COUNT()` y `SUM()`.
  - Filtros que aplican `CAST()` a un atributo. Por ejemplo, `CAST(stringColumn as INT) = 1`.
  - Filtros con un atributo que es un objeto o es complejo. Por ejemplo, `intArray[1] = 1`, `objectColumn.objectNumber = 1`.
  - Filtros cuyo valor no es un valor literal. Por ejemplo, `intColumn1 = intColumn2`
  - Solo se admiten los tipos de datos compatibles con S3 Select con las limitaciones que se especifican.

Especificación de S3 Select en el código

Los siguientes ejemplos muestran cómo especificar S3 Select para CSV mediante Scala, SQL, R y PySpark. Puede utilizar S3 Select para JSON de la misma manera. Para obtener una lista de las opciones, sus valores predeterminados y sus limitaciones, consulte Opciones.

PySpark

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional
  .load("s3://path/to/my/datafiles")
```

R

```
read.df("s3://path/to/my/datafiles", "s3selectCSV", schema, header = "true",
 delimiter = "\t")
```

Scala

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional. Examples:
  // .options(Map("quote" -> "\'", "header" -> "true")) or
  // .option("quote", "\'").option("header", "true")
  .load("s3://path/to/my/datafiles")
```

SQL

```
CREATE TEMPORARY VIEW MyView (number INT, name STRING) USING s3selectCSV OPTIONS
 (path "s3://path/to/my/datafiles", header "true", delimiter "\t")
```

Opciones

Las siguientes opciones están disponibles cuando se utilizan `s3selectCSV` y `s3selectJSON`. Si no se especifican, se usarán los valores predeterminados.

Opciones con S3selectCSV

Opción	Predeterminado/a	Uso
`compression`	`"none"`	Indica si se utiliza la compresión. `"gzip"` es la única opción admitida además de `"none"`.
`delimiter`	`","`	Especifica el delimitador de campo.
`quote`	`'\"'`	Especifica el carácter de comillas. No es posible

Opción	Predeterminado/a	Uso
		especificar una cadena vacía. Si lo hace, se mostrará un error que indica que el XML no tiene el formato correcto.
`escape`	`'\\'`	Especifica el carácter de escape.
`header`	`"false"`	`"false"` especifica que no existe ningún encabezado. `"true"` especifica que hay un encabezado en la primera línea. Solo se admiten los encabezados en la primera línea, y no se admiten las líneas vacías antes del encabezado.
comentario	`"#"`	Especifica el carácter de comentario. El indicador de comentario no se puede deshabilitar. En otras palabras, no se admite el valor `\u0000`.
`nullValue`	`""`	

Opciones con S3selectJSON

Opción	Predeterminado/a	Uso
`compression`	`"none"`	Indica si se utiliza la compresión. `"gzip"` es la única opción admitida además de `"none"`.

Opción	Predeterminado/a	Uso
`multiline`	"false"	`"false"` especifica que el JSON está en el formato `LINES` de S3 Select, lo que significa que cada línea de los datos de entrada contiene un único objeto JSON. `"true"` especifica que el JSON está en el formato `DOCUMENT` de S3 Select, lo que significa que un objeto JSON pueden abarcar varias líneas en los datos de entrada.

Uso del confirmador optimizado para S3 de EMRFS

El confirmador optimizado para S3 para EMRFS es una OutputCommitterimplementación alternativa que está optimizada para escribir archivos en Amazon S3 cuando se usa EMRFS. El confirmador optimizado para S3 de EMRFS mejora el rendimiento de las aplicaciones al evitar que se lleven a cabo operaciones de cambio de nombre y de listado en Amazon S3 durante las fases de confirmación de tareas y de trabajos. El confirmador está disponible con la versión 5.19.0 y posteriores de Amazon EMR y está habilitado de forma predeterminada con la versión 5.20.0 y posteriores de Amazon EMR. El confirmador se utiliza para los trabajos de Spark que utilizan Spark o Datasets. DataFrames A partir de Amazon EMR 6.4.0, este confirmador se puede utilizar para todos los formatos habituales, incluidos los formatos Parquet, ORC y basados en texto (incluidos CSV y JSON). En el caso de las versiones anteriores a Amazon EMR 6.4.0, solo se admite el formato Parquet. Existen circunstancias bajo las cuales no se utiliza el confirmador. Para obtener más información, consulte Requisitos del confirmador optimizado para S3 de EMRFS.

Temas

- Requisitos del confirmador optimizado para S3 de EMRFS
- El confirmador optimizado para S3 de EMRFS y las cargas multiparte
- Consideraciones sobre el ajuste de trabajos
- Habilitación del confirmador optimizado para S3 de EMRFS para Amazon EMR 5.19.0

Requisitos del confirmador optimizado para S3 de EMRFS

El confirmador optimizado para S3 de EMRFS se utiliza cuando se cumplen las siguientes condiciones:

- Ejecuta trabajos de Spark que utilizan Spark o Datasets para escribir archivos en Amazon S3. DataFrames A partir de Amazon EMR 6.4.0, este confirmador se puede utilizar para todos los formatos habituales, incluidos los formatos Parquet, ORC y basados en texto (incluidos CSV y JSON). En el caso de las versiones anteriores a Amazon EMR 6.4.0, solo se admite el formato Parquet.
- Las cargas multiparte están habilitadas en Amazon EMR. Esta es la opción predeterminada. Para obtener más información, consulte El confirmador optimizado para S3 de EMRFS y las cargas multiparte.
- Se utiliza la compatibilidad de formatos de archivos integrados de Spark. La compatibilidad de formatos de archivos integrados se utiliza en las siguientes circunstancias:
  - En el caso de las tablas del almacén de metadatos de Hive, cuando `spark.sql.hive.convertMetastoreParquet` se establece en `true` para tablas Parquet o `spark.sql.hive.convertMetastoreOrc` se establece en `true` para tablas ORC con Amazon EMR 6.4.0 o posteriores. Esta es la configuración predeterminada.
  - Cuando los trabajos escriben orígenes de datos o tablas de formatos de archivos, por ejemplo, la tabla de destino se crea con la cláusula `USING parquet`.
  - Cuando los trabajos escriben en tablas de Parquet de metaalmacenes de Hive no particionadas. La compatibilidad con Parquet integrada de Spark no admite tablas de Hive particionadas, se trata de una limitación conocida. Para obtener más información, consulte la conversión de tablas de Hive Metastore Parquet en la Guía de conjuntos de datos DataFrames y Spark de Apache.
- Las operaciones de trabajos de Spark que escriben en una ubicación de partición predeterminada (por ejemplo, `${table_location}/k1=v1/k2=v2/`) utilizan el confirmador. El confirmador no se utiliza si una operación de trabajo escribe en una ubicación de partición personalizada, por ejemplo, si una ubicación de partición personalizada se establece con el comando `ALTER TABLE SQL`.
- Se deben utilizar los valores siguientes para Spark:
  - La propiedad `spark.sql.parquet.fs.optimized.committer.optimization-enabled` debe establecerse en `true`. Esta es la configuración predeterminada con Amazon EMR 5.20.0 y versiones posteriores. Con Amazon EMR 5.19.0, el valor predeterminado es `false`. Para

obtener más información acerca de cómo configurar este valor, consulte Habilitación del confirmador optimizado para S3 de EMRFS para Amazon EMR 5.19.0.

- Si se escribe en tablas del metaalmacén de Hive no particionadas, solo se admiten los formatos de archivo Parquet y Orc. `spark.sql.hive.convertMetastoreParquet` debe establecerse en `true` si se escribe en tablas Parquet no particionadas del metaalmacén de Hive. `spark.sql.hive.convertMetastoreOrc` debe establecerse en `true` si se escribe en tablas Orc no particionadas del metaalmacén de Hive. Esta es la configuración predeterminada.
- `spark.sql.parquet.output.committer.class` se debe establecer en `com.amazon.emr.committer.EmrOptimizedSparkSqlParquetOutputCommitter`. Este es el valor predeterminado.
- `spark.sql.sources.commitProtocolClass` debe establecerse en `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` o `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol`. `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol` es la configuración predeterminada para la versión 5.30.0 y posteriores de la serie 5.x de Amazon EMR y para la versión 6.2.0 y posteriores de la serie 6.x de Amazon EMR. `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` es la configuración predeterminada para las versiones anteriores de Amazon EMR.
- Si los trabajos de Spark sobrescriben conjuntos de datos de Parquet particionados con columnas de partición, entonces la opción de escritura `partitionOverwriteMode` y `spark.sql.sources.partitionOverwriteMode` se deben establecer en `static`. Este es el valor predeterminado.

> **Note**
>
> La opción de escritura `partitionOverwriteMode` se introdujo en Spark 2.4.0. En el caso de la versión 2.3.2 de Spark, incluida con la versión 5.19.0 de Amazon EMR, establezca la propiedad `spark.sql.sources.partitionOverwriteMode`.

Ocasiones en las que no se utiliza el confirmador optimizado para S3 de EMRFS

Por lo general, el confirmador optimizado para S3 de EMRFS no se utiliza en las siguientes situaciones.

Situación	Por qué no se utiliza el confirmador
Cuando escribe en HDFS	El confirmador solo admite la escritura en Amazon S3 mediante EMRFS.
Cuando utiliza el sistema de archivos S3A	El confirmador solo admite EMRFS.
Cuando utilizas nuestra API RDD MapReduce de Spark	El confirmador solo admite el uso de SparkSQL o DataFrame Dataset. APIs

En los siguientes ejemplos se muestran algunas situaciones más que evitan el uso del confirmador optimizado para S3 de EMRFS en su conjunto (el primer ejemplo) y en parte (el segundo ejemplo).

Example – Modo de sobrescritura de partición dinámica

En el siguiente ejemplo de Scala se indica a Spark que utilice un algoritmo de confirmación diferente, lo que impide por completo el uso del confirmador optimizado para S3 de EMRFS. El código establece la propiedad `partitionOverwriteMode` en `dynamic` para sobrescribir solo las particiones en las que escriba datos. A continuación, las columnas de particiones dinámicas se especifican mediante `partitionBy` y el modo de escritura se establece en `overwrite`.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")
  .option("partitionOverwriteMode", "dynamic")
  .partitionBy("dt")
  .parquet("s3://amzn-s3-demo-bucket1/output")
```

Debe establecer las tres configuraciones para evitar utilizar el confirmador optimizado para S3 de EMRFS. Al hacerlo, Spark ejecuta un algoritmo de confirmación diferente que se especifica en el protocolo de confirmación de Spark. En el caso de las versiones 5.x de Amazon EMR anteriores a la 5.30.0 y las versiones 6.x de Amazon EMR anteriores a la 6.2.0, el protocolo de confirmación utiliza el directorio provisional de Spark, que es un directorio temporal creado en la ubicación de salida que comienza por `.spark-staging`. El algoritmo cambia de forma secuencial el nombre de los directorios de particiones, lo que puede afectar negativamente al rendimiento. Para obtener más información sobre la versión 5.30.0 y posteriores y 6.2.0 y posteriores de Amazon EMR, consulte Uso del protocolo de confirmación optimizado para S3 de EMRFS.

El algoritmo de Spark 2.4.0 sigue estos pasos:

1. La tarea intenta escribir su salida en los directorios de particiones en el directorio provisional de Spark (por ejemplo, `${outputLocation}/spark-staging-${jobID}/k1=v1/k2=v2/`).
2. Para cada partición escrita, la tarea intenta llevar un seguimiento de las rutas de partición relativas (por ejemplo, `k1=v1/k2=v2`).
3. Cuando una tarea se completa de forma satisfactoria, proporciona al controlador todas las rutas de partición relativa cuyo seguimiento ha realizado.
4. Una vez completadas todas las tareas, la fase de confirmación del trabajo recopila todos los directorios de partición que los intentos de tarea correctos escribieron en el directorio de ensayo de Spark. Spark cambia el nombre de estos directorios de forma secuencial a su ubicación de salida final utilizando operaciones de cambio de nombre de árbol de directorio.
5. El directorio de ensayo se elimina antes de que se complete la fase de confirmación de trabajo.

Example – Ubicación de partición personalizada

En este ejemplo, el código Scala se inserta en dos particiones. Una partición tiene una ubicación de partición personalizada. La otra partición utiliza la ubicación de partición predeterminada. El confirmador optimizado para S3 de EMRFS se utiliza solo para escribir la salida de tarea en la partición que utiliza la ubicación de partición predeterminada.

```
val table = "dataset"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
```

```
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .write.insertInto(table)
```

El código de Scala crea los siguientes objetos de Amazon S3:

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

Al escribir en particiones de ubicaciones personalizadas, Spark utiliza un algoritmo de confirmación similar al del ejemplo anterior, que se detalla a continuación. Como ocurre en el ejemplo anterior, el algoritmo da como resultado cambios de nombre secuenciales, lo que puede afectar negativamente al rendimiento.

1. Al escribir la salida en una partición de una ubicación personalizada, las tareas escriben en un archivo del directorio de ensayo de Spark, que se crea en la ubicación de salida final. El nombre del archivo incluye un UUID aleatorio para proteger contra las colisiones de archivo. La tarea intenta mantener un seguimiento de cada archivo junto con la ruta de salida deseada final.
2. Cuando una tarea se completa de forma satisfactoria, proporciona al controlador los archivos y las rutas de salida deseada final.
3. Una vez completadas todas las tareas, la fase de confirmación de trabajo cambia el nombre de forma secuencial de todos los archivos que se escribieron para particiones en ubicaciones personalizadas en sus rutas de salida finales.
4. El directorio de ensayo se elimina antes de que se complete la fase de confirmación de trabajo.

El confirmador optimizado para S3 de EMRFS y las cargas multiparte

Para utilizar el confirmador optimizado para S3 de EMRFS, debe habilitar las cargas multiparte para Amazon EMR. Las cargas multiparte están habilitadas de forma predeterminada. Puede volver

a habilitarlas si es necesario. Para obtener más información, consulte Configuración de carga multiparte para Amazon S3 en la Guía de administración de Amazon EMR.

El confirmador optimizado para S3 de EMRFS utiliza las características similares a las transacciones de las cargas multiparte para garantizar que los archivos escritos por los intentos de tareas solo aparecen en la ubicación de salida del trabajo al confirmar las tareas. Al utilizar las cargas multiparte de esta forma, el confirmador mejora el rendimiento de las confirmaciones de tareas en comparación con la versión 2 del algoritmo predeterminado. FileOutputCommitter Cuando se utiliza el confirmador optimizado para S3 de EMRFS, existen algunas diferencias clave que lo distinguen del comportamiento tradicional de las cargas multiparte y que hay que tener en cuenta:

- Las cargas multiparte siempre se llevan a cabo independientemente del tamaño del archivo. Esto difiere del comportamiento predeterminado de EMRFS, donde la propiedad `fs.s3n.multipart.uploads.split.size` controla el tamaño del archivo en el que se activan las cargas multiparte.
- Las cargas multiparte se dejan en un estado incompleto durante un periodo de tiempo más prolongado hasta que la tarea se confirma o se anula. Esto difiere del comportamiento predeterminado de EMRFS, donde una carga multiparte se completa cuando una tarea termina de escribir un determinado archivo.

Debido a estas diferencias, si la JVM de un ejecutor de Spark se bloquea o se termina mientras hay tareas ejecutándose y escribiendo datos en Amazon S3, es más probable que queden cargas multiparte incompletas sin finalizar. Por este motivo, cuando utilice el confirmador optimizado para S3 de EMRFS, asegúrese de seguir las prácticas recomendadas para la administración de cargas multiparte con errores. Para obtener más información, consulte Prácticas recomendadas para trabajar con buckets de Amazon S3 en la Guía de administración de Amazon EMR.

Consideraciones sobre el ajuste de trabajos

El confirmador optimizado para S3 de EMRFS consume una pequeña cantidad de memoria por cada archivo escrito por un intento de tarea hasta que la tarea se confirma o se anula. En la mayoría de los trabajos, la cantidad de memoria consumida es insignificante. En los trabajos que tienen tareas de ejecución prolongada que escriben un gran número de archivos, la memoria que consume el confirmador puede ser apreciable y requerir algunos ajustes de la memoria asignada a los ejecutores de Spark. Puede ajustar la memoria del ejecutor mediante la propiedad `spark.executor.memory`. Como pauta general, una sola tarea que escribe 100 000 archivos, normalmente requerirá 100 MB

de memoria adicionales. Para obtener más información, consulte Application properties en la documentación de configuración de Apache Spark.

Habilitación del confirmador optimizado para S3 de EMRFS para Amazon EMR 5.19.0

Si utiliza Amazon EMR 5.19.0, puede establecer manualmente la propiedad `spark.sql.parquet.fs.optimized.committer.optimization-enabled` en `true` al crear un clúster o desde Spark si está utilizando Amazon EMR.

Habilitación del confirmador optimizado para S3 de EMRFS al crear un clúster

Utilice la clasificación de configuración `spark-defaults` para establecer la propiedad `spark.sql.parquet.fs.optimized.committer.optimization-enabled` en `true`. Para obtener más información, consulte Configuración de aplicaciones.

Habilitación del confirmador optimizado para S3 de EMRFS desde Spark

Puede establecer la propiedad `spark.sql.parquet.fs.optimized.committer.optimization-enabled` en `true` codificándola de forma rígida en un `SparkConf`, pasándola como un parámetro `--conf` en el shell de Spark o las herramientas `spark-submit` y `spark-sql`, o en `conf/spark-defaults.conf`. Para obtener más información, consulte Spark configuration en la documentación de Apache Spark.

En el siguiente ejemplo, se muestra cómo activar el confirmador mientras se ejecuta un comando spark-sql.

```
spark-sql \
  --conf spark.sql.parquet.fs.optimized.committer.optimization-enabled=true \
  -e "INSERT OVERWRITE TABLE target_table SELECT * FROM source_table;"
```

Uso del protocolo de confirmación optimizado para S3 de EMRFS

El protocolo de confirmación optimizado para S3 para EMRFS es una FileCommitProtocolimplementación alternativa que está optimizada para escribir archivos con la sobrescritura dinámica de particiones de Spark en Amazon S3 cuando se usa EMRFS. El protocolo mejora el rendimiento de las aplicaciones al evitar las operaciones de cambio de nombre en Amazon S3 durante la fase de confirmación de los trabajos de sobrescritura de particiones dinámicas de Spark.

Tenga en cuenta que el Uso del confirmador optimizado para S3 de EMRFS también mejora el rendimiento al evitar las operaciones de cambio de nombre. Sin embargo, no funciona para los casos

de sobrescritura de particiones dinámicas, mientras que las mejoras del protocolo de confirmación solo se dirigen a los casos de sobrescritura de particiones dinámicas.

El protocolo de confirmación está disponible con la versión 5.30.0 y posteriores y 6.2.0 y posteriores de Amazon EMR y está habilitado de forma predeterminada. Amazon EMR agregó una mejora de paralelismo a partir de la versión 5.31.0. El protocolo se utiliza para los trabajos de Spark que utilizan Spark o Datasets. DataFrames Existen circunstancias en las cuales no se utiliza el protocolo de confirmación. Para obtener más información, consulte Requisitos del protocolo de confirmación optimizado para S3 de EMRFS.

Temas

- Requisitos del protocolo de confirmación optimizado para S3 de EMRFS
- El protocolo de confirmación optimizado para S3 de EMRFS y las cargas multiparte
- Consideraciones sobre el ajuste de trabajos

Requisitos del protocolo de confirmación optimizado para S3 de EMRFS

El protocolo de confirmación optimizado para S3 de EMRFS se utiliza cuando se cumplen las siguientes condiciones:

- Ejecutas trabajos de Spark que usan Spark o Datasets para sobrescribir tablas particionadas. DataFrames
- Ejecuta trabajos de Spark cuyo modo de sobrescritura de particiones es `dynamic`.
- Las cargas multiparte están habilitadas en Amazon EMR. Esta es la opción predeterminada. Para obtener más información, consulte El protocolo de confirmación optimizado para S3 de EMRFS y las cargas multiparte.
- La caché del sistema de archivos para EMRFS está habilitada. Esta es la opción predeterminada. Compruebe que la configuración `fs.s3.impl.disable.cache` esté establecida en `false`.
- Se utiliza la compatibilidad de orígenes de datos integrados de Spark. La compatibilidad de orígenes de datos integrados se utiliza en las siguientes circunstancias:
  - Cuando los trabajos escriben en orígenes de datos o tablas integrados.
  - Cuando los trabajos escriben en tablas Parquet del metaalmacén de Hive. Esto sucede cuando `spark.sql.hive.convertInsertingPartitionedTable` y `spark.sql.hive.convertMetastoreParquet` se establecen en true. Esta es la configuración predeterminada.

  - Cuando los trabajos escriben en tablas ORC del metaalmacén de Hive. Esto sucede cuando `spark.sql.hive.convertInsertingPartitionedTable` y `spark.sql.hive.convertMetastoreOrc` se establecen en `true`. Esta es la configuración predeterminada.

- Las operaciones de trabajos de Spark que escriben en una ubicación de partición predeterminada (por ejemplo, `${table_location}/k1=v1/k2=v2/`) utilizan el protocolo de confirmación. El protocolo no se utiliza si una operación de trabajo escribe en una ubicación de partición personalizada, por ejemplo, si una ubicación de partición personalizada se establece con el comando `ALTER TABLE SQL`.
- Se deben utilizar los valores siguientes para Spark:
  - `spark.sql.sources.commitProtocolClass` se debe establecer en `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol`. Esta es la configuración predeterminada para las versiones 5.30.0 y posteriores y 6.2.0 y posteriores de Amazon EMR.
  - La opción de escritura `partitionOverwriteMode` o `spark.sql.sources.partitionOverwriteMode` debe estar establecida en `dynamic`. El ajuste predeterminado es `static`.

    > **Note**
    >
    > La opción de escritura `partitionOverwriteMode` se introdujo en Spark 2.4.0. En el caso de la versión 2.3.2 de Spark, incluida con la versión 5.19.0 de Amazon EMR, establezca la propiedad `spark.sql.sources.partitionOverwriteMode`.

  - Si los trabajos de Spark sobrescriben en la tabla Parquet del metaalmacén de Hive, `spark.sql.hive.convertMetastoreParquet`, `spark.sql.hive.convertInsertingPartitionedTable` y `spark.sql.hive.convertMetastore.partitionOverwriteMode` debe establecerse en `true`. Esta es la configuración predeterminada.
  - Si los trabajos de Spark sobrescriben en la tabla ORC del metaalmacén de Hive, `spark.sql.hive.convertMetastoreOrc`, `spark.sql.hive.convertInsertingPartitionedTable` y `spark.sql.hive.convertMetastore.partitionOverwriteMode` debe establecerse en `true`. Esta es la configuración predeterminada.

Example – Modo de sobrescritura de partición dinámica

En este ejemplo de Scala, se activa la optimización. En primer lugar, establece la propiedad `partitionOverwriteMode` en `dynamic`. Esto solo sobrescribe las particiones en las que escribe datos. A continuación, especifica las columnas de particiones dinámicas mediante `partitionBy` y establece el modo de escritura en `overwrite`.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")                 // "overwrite" instead of "insert"
  .option("partitionOverwriteMode", "dynamic")  // "dynamic" instead of "static"
  .partitionBy("dt")                            // partitioned data instead of
 unpartitioned data
  .parquet("s3://amzn-s3-demo-bucket1/output")    // "s3://" to use Amazon EMR file
 system, instead of "s3a://" or "hdfs://"
```

Cuando no se utiliza el protocolo de confirmación optimizado para S3 de EMRFS

Por lo general, el protocolo de confirmación optimizado para EMRFS para S3 funciona igual que el protocolo de confirmación de Spark predeterminado de código abierto,. `org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol` La optimización no se producirá en las siguientes situaciones.

Situación	Por qué no se utiliza el protocolo de confirmac ión
Cuando escribe en HDFS	El protocolo de confirmación solo admite la escritura en Amazon S3 mediante EMRFS.
Cuando utiliza el sistema de archivos S3A	El protocolo de confirmación solo admite EMRFS.
Cuando utilizas nuestra API RDD de MapReduce Spark	El protocolo de confirmación solo admite el uso de SparkSQL o DataFrame Dataset. APIs
Cuando no se activa la sobrescritura de particiones dinámicas	El protocolo de confirmación solo optimiza los casos de sobrescritura de particiones dinámicas. Para los demás casos, consulte

Situación	Por qué no se utiliza el protocolo de confirmación
	Uso del confirmador optimizado para S3 de EMRFS.

En los siguientes ejemplos de Scala se muestran algunas situaciones adicionales en las que el protocolo de confirmación optimizado para S3 de EMRFS delega en `SQLHadoopMapReduceCommitProtocol`.

Example – Modo de sobrescritura de particiones dinámica con ubicación de partición personalizada

En este ejemplo, los programas de Scala sobrescriben dos particiones en el modo de sobrescritura de particiones dinámicas. Una partición tiene una ubicación de partición personalizada. La otra partición utiliza la ubicación de partición predeterminada. El protocolo de confirmación optimizado para S3 de EMRFS solo mejora la partición que utiliza la ubicación de partición predeterminada.

```
val table = "dataset"
val inputView = "tempView"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
```

```
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .createTempView(inputView)

// Set partition overwrite mode to 'dynamic'
spark.sql(s"SET spark.sql.sources.partitionOverwriteMode=dynamic")

spark.sql(s"INSERT OVERWRITE TABLE $table SELECT * FROM $inputView")
```

El código de Scala crea los siguientes objetos de Amazon S3:

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

> **Note**
>
> Escribir en ubicaciones de particiones personalizadas en versiones anteriores de Spark puede provocar la pérdida de datos. En este ejemplo, la partición `dt='2019-01-28'` se perdería. Para obtener más información, consulte SPARK-35106. Esto se corrigió en la versión 5.33.0 y posteriores de Amazon EMR, excluidas las 6.0.x y 6.1.x.

Al escribir en particiones de ubicaciones personalizadas, Spark utiliza un algoritmo de confirmación similar al del ejemplo anterior, que se detalla a continuación. Como ocurre en el ejemplo anterior, el algoritmo da como resultado cambios de nombre secuenciales, lo que puede afectar negativamente al rendimiento.

El algoritmo de Spark 2.4.0 sigue estos pasos:

1. Al escribir la salida en una partición de una ubicación personalizada, las tareas escriben en un archivo del directorio de ensayo de Spark, que se crea en la ubicación de salida final. El nombre del archivo incluye un UUID aleatorio para proteger contra las colisiones de archivo. La tarea intenta mantener un seguimiento de cada archivo junto con la ruta de salida deseada final.
2. Cuando una tarea se completa de forma satisfactoria, proporciona al controlador los archivos y las rutas de salida deseada final.

3. Una vez completadas todas las tareas, la fase de confirmación de trabajo cambia el nombre de forma secuencial de todos los archivos que se escribieron para particiones en ubicaciones personalizadas en sus rutas de salida finales.
4. El directorio de ensayo se elimina antes de que se complete la fase de confirmación de trabajo.

El protocolo de confirmación optimizado para S3 de EMRFS y las cargas multiparte

Para utilizar la optimización de la sobrescritura de particiones dinámicas en el protocolo de confirmación optimizado para S3 de EMRFS, las cargas multiparte deben estar habilitadas en Amazon EMR. Las cargas multiparte están habilitadas de forma predeterminada. Puede volver a habilitarlas si es necesario. Para obtener más información, consulte Configuración de carga multiparte para Amazon S3 en la Guía de administración de Amazon EMR.

Durante la sobrescritura de particiones dinámicas, el protocolo de confirmación optimizado para S3 de EMRFS utiliza las características similares a las transacciones de las cargas multiparte para garantizar que los archivos escritos por los intentos de tareas solo aparecen en la ubicación de salida del trabajo al confirmar las tareas. Al utilizar las cargas multiparte de este modo, el protocolo de confirmación mejora el rendimiento de las confirmaciones de trabajos en comparación con la opción predeterminada `SQLHadoopMapReduceCommitProtocol`. Cuando se utiliza el protocolo de confirmación optimizado para S3 de EMRFS, hay algunas diferencias clave que lo distinguen del comportamiento tradicional de las cargas multiparte y que hay que tener en cuenta:

- Las cargas multiparte siempre se llevan a cabo independientemente del tamaño del archivo. Esto difiere del comportamiento predeterminado de EMRFS, donde la propiedad `fs.s3n.multipart.uploads.split.size` controla el tamaño del archivo en el que se activan las cargas multiparte.
- Las cargas multiparte se dejan en un estado incompleto durante un periodo de tiempo más prolongado hasta que la tarea se confirma o se anula. Esto difiere del comportamiento predeterminado de EMRFS, donde una carga multiparte se completa cuando una tarea termina de escribir un determinado archivo.

Debido a estas diferencias, si la JVM de un ejecutor de Spark se bloquea o se termina mientras hay tareas ejecutándose y escribiendo datos en Amazon S3, o la JVM de un controlador de Spark se bloquea o se termina durante la ejecución de un trabajo, es más probable que queden cargas multiparte incompletas sin finalizar. Por este motivo, cuando utilice el protocolo de confirmación optimizado para S3 de EMRFS, asegúrese de seguir las prácticas recomendadas para la

administración de cargas multiparte con errores. Para obtener más información, consulte Prácticas recomendadas para trabajar con buckets de Amazon S3 en la Guía de administración de Amazon EMR.

Consideraciones sobre el ajuste de trabajos

En los ejecutores de Spark, el protocolo de confirmación optimizado para S3 de EMRFS consume una pequeña cantidad de memoria por cada archivo escrito por un intento de tarea hasta que la tarea se confirma o se anula. En la mayoría de los trabajos, la cantidad de memoria consumida es insignificante.

En los controladores de Spark, el protocolo de confirmación optimizado para S3 de EMRFS requiere memoria para almacenar la información de los metadatos de cada archivo confirmado hasta que el trabajo se confirma o se anula. En la mayoría de los trabajos, la configuración de memoria predeterminada del controlador de Spark es insignificante.

En los trabajos que tienen tareas de ejecución prolongada y que escriben un gran número de archivos, la memoria que consume el protocolo de confirmación puede ser apreciable y requerir algunos ajustes de la memoria asignada a Spark, especialmente a los ejecutores de Spark. Puede ajustar la memoria mediante la propiedad `spark.driver.memory` para los controladores de Spark y la propiedad `spark.executor.memory` para los ejecutores de Spark. Como pauta general, una sola tarea que escribe 100 000 archivos, normalmente requerirá 100 MB de memoria adicionales. Para obtener más información, consulte Application properties en la documentación de configuración de Apache Spark.

Reintento de solicitudes de Amazon S3 con EMRFS

En este tema se proporciona información sobre las estrategias de reintentos que puede utilizar al realizar solicitudes a Amazon S3 con EMRFS. Cuando la tasa de solicitudes aumenta, S3 intenta escalar para dar cabida a la nueva tasa. Durante este proceso, S3 puede limitar las solicitudes y devolver un error `503 Slow Down`. Para mejorar la tasa de éxito de las solicitudes de S3, puede configurar las propiedades de la configuración `emrfs-site` para ajustar la estrategia de reintentos.

Puede ajustar la estrategia de reintentos de las siguientes maneras.

- Aumente el límite máximo de reintentos para la estrategia de reintentos de retroceso exponencial predeterminada.
- Habilite y configure la estrategia de reintentos de aumento aditivo/disminución multiplicativa (AIMD). AIMD es compatible con la versión 6.4.0 y posteriores de Amazon EMR.

Uso de la estrategia de retroceso exponencial predeterminada

De forma predeterminada, EMRFS utiliza una estrategia de retroceso exponencial para reintentar las solicitudes de Amazon S3. El límite de reintentos de EMRFS predeterminado es de 15. Para evitar un error `503 Slow Down` de S3, puede aumentar el límite de reintentos al crear un clúster nuevo, en un clúster en ejecución o durante el tiempo de ejecución de la aplicación.

Para aumentar el límite de reintentos, debe cambiar el valor de `fs.s3.maxRetries` de la configuración `emrfs-site`. En el siguiente ejemplo de configuración se establece `fs.s3.maxRetries` en un valor personalizado de 30.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.maxRetries": "30"
      }
    }
]
```

Para obtener información acerca del uso de objetos de configuración, consulte Configuración de aplicaciones.

Uso de la estrategia de reintentos de AIMD

Con la versión 6.4.0 y posteriores de Amazon EMR, EMRFS admite una estrategia de reintentos alternativa basada en un modelo de aumento aditivo/disminución multiplicativa (AIMD). La estrategia de reintentos de AIMD resulta especialmente útil cuando se trabaja con clústeres grandes de Amazon EMR.

AIMD calcula una tasa de solicitudes personalizada a partir de los datos sobre las solicitudes recientes que se han realizado correctamente. Esta estrategia reduce el número de solicitudes limitadas y el total de intentos necesarios por solicitud.

Para habilitar la estrategia de reintentos de AIMD, debe establecer la propiedad `fs.s3.aimd.enabled` en `true` en la configuración `emrfs-site`, como en el siguiente ejemplo.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
```

```
        "fs.s3.aimd.enabled": "true"
      }
    }
]
```

Para obtener información acerca del uso de objetos de configuración, consulte Configuración de aplicaciones.

Configuración avanzada de reintentos de AIMD

Puede configurar las propiedades que se muestran en la siguiente tabla para refinar el comportamiento de los reintentos al utilizar la estrategia de reintentos de AIMD. Para la mayoría de los casos de uso, le recomendamos que utilice los valores predeterminados.

Propiedades avanzadas de la estrategia de reintentos de AIMD

Propiedad	Valor predeterminado	Descripción
`fs.s3.aimd.increaseIncrement`	0.1	Controla la rapidez con la que aumenta la tasa de solicitudes cuando las solicitudes consecutivas se realizan correctamente.
`fs.s3.aimd.reductionFactor`	2	Controla la rapidez con la que disminuye la tasa de solicitudes cuando Amazon S3 devuelve una respuesta 503. El factor predeterminado de 2 reduce la tasa de solicitudes a la mitad.
`fs.s3.aimd.minRate`	0.1	Establece el límite inferior de la tasa de solicitudes cuando las solicitudes sufren una limitación constante por parte de S3.
`fs.s3.aimd.initialRate`	5500	Establece la tasa de solicitudes inicial, que luego cambia

Propiedad	Valor predeterminado	Descripción
		según los valores que especifique para `fs.s3.aimd.increaseIncrement` y `fs.s3.aimd.reductionFactor` . La tasa inicial también se usa para las solicitudes GET y se escala de forma proporcional (3500/5500) para las solicitudes PUT.
`fs.s3.aimd.adjustWindow`	2	Controla la frecuencia con la que se ajusta la tasa de solicitudes, medida en número de respuestas.
`fs.s3.aimd.maxAttempts`	100	Establece el número máximo de intentos para intentar una solicitud.

Adición de un paso de Spark

Puede utilizar los pasos de Amazon EMR para enviar trabajo al marco de Spark instalado en un clúster de EMR. Para obtener más información, consulte Pasos en la Guía de administración de Amazon EMR. En la consola y la interfaz de línea de comandos, puede hacerlo a través de un paso de la aplicación Spark, que ejecuta el script `spark-submit` como un paso en su nombre. Con la API, utilice un paso para invocar `spark-submit` utilizando `command-runner.jar`.

Para obtener más información acerca de cómo enviar aplicaciones a Spark, consulte Submitting applications en la documentación de Apache Spark.

Para enviar un paso de Spark utilizando la consola

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En Cluster List (Lista de clústeres), elija el nombre del clúster.

3. Desplácese hasta la sección Steps (Pasos) y amplíela; a continuación, elija Add step (Añadir paso).
4. En el cuadro de diálogo Add Step (Añadir paso):
   - En Step type (Tipo de paso), elija Spark application (Aplicación de Spark).
   - En Name (Nombre), acepte el nombre predeterminado (Aplicación de Spark) o escriba un nombre nuevo.
   - En Deploy mode (Modo de implementación), elija el modo Client (Cliente) o Cluster (Clúster). El modo de cliente lanza el programa de controlador en la instancia principal del clúster, mientras que el modo de clúster lanza el programa de controlador en el clúster. En el modo de cliente, la salida del registro del controlador aparece en los registros de pasos, mientras que en el modo de clúster, la salida del registro del controlador aparece en los registros del primer contenedor de YARN. Para obtener más información, consulte Cluster mode overview en la documentación de Apache Spark.
   - Especifique el valor deseado de Spark-submit options (Opciones de spark-submit). Para obtener más información sobre las opciones de `spark-submit`, consulte Launching applications with spark-submit.
   - En Application location (Ubicación de la aplicación), especifique la ruta local o el URI de S3 de la aplicación.
   - En Arguments (Argumentos), deje el campo en blanco.
   - En Action on failure (Acción sobre el error), acepte la opción predeterminada, Continue (Continuar).
5. Elija Agregar. El paso aparece en la consola con el estado Pending (Pendiente).
6. El estado del paso cambia de Pending (Pendiente) a Running (En ejecución) y a Completed (Completado) a medida que se ejecuta. Para actualizar el estado, elija el icono Refresh (Actualizar) situado encima de la columna Actions (Acciones).
7. Los resultados del paso se encuentran en la página Detalles del clúster de la consola de Amazon EMR junto a su paso en Archivos de registro si ha configurado el registro. Si lo desea, puede encontrar información sobre el paso en el bucket de registro que configuró al lanzar el clúster.

Para enviar trabajos a Spark utilizando el AWS CLI

Envíe un paso al crear el clúster o utilice el subcomando `aws emr add-steps` en un clúster existente.

1. Use `create-cluster` como se muestra en el ejemplo siguiente.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 --
applications Name=Spark \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10] --use-default-roles
```

También puede usar `command-runner.jar` como se muestra en el siguiente ejemplo.

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 \
--applications Name=Spark --ec2-attributes KeyName=myKey --instance-type m5.xlarge
 --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Spark Program",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[spark-example,SparkPi,10] --use-default-
roles
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

2. Como alternativa, añada los pasos a un clúster ya en ejecución. Utilice `add-steps`.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps
 Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10]
```

También puede usar `command-runner.jar` como se muestra en el siguiente ejemplo.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Spark
 Program",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[spark-
example,SparkPi,10]
```

Para enviar trabajo a Spark mediante el SDK para Java

1. En el siguiente ejemplo se muestra cómo agregar un paso a un clúster con Spark mediante Java.

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);

StepFactory stepFactory = new StepFactory();
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig sparkStepConf = new HadoopJarStepConfig()
   .withJar("command-runner.jar")
   .withArgs("spark-submit","--executor-memory","1g","--
class","org.apache.spark.examples.SparkPi","/usr/lib/spark/examples/jars/spark-
examples.jar","10");

StepConfig sparkStep = new StepConfig()
   .withName("Spark Step")
   .withActionOnFailure("CONTINUE")
   .withHadoopJarStep(sparkStepConf);

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

2. Vea los resultados del paso examinando los registros del paso. Para ello, AWS Management Console si has activado el registro, selecciona Pasos, selecciona tu paso y, en el caso de los archivos de registro, elige una de las dos `stdout` opciones`stderr`. Para ver los registros disponibles, elija View Logs (Ver registros).

Anulación de las opciones de configuración predeterminadas de Spark

Es posible que desee anular los valores de configuración predeterminados de Spark por aplicación. Puede hacerlo cuando envíe aplicaciones utilizando un paso, que básicamente transfiere opciones a `spark-submit`. Por ejemplo, es posible que desee cambiar la memoria asignada a un proceso de ejecutor cambiando `spark.executor.memory`. Suministraría el modificador `--executor-memory` con un argumento como el siguiente:

```
spark-submit --executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/
spark/examples/jars/spark-examples.jar 10
```

Del mismo modo, puede ajustar `--executor-cores` y `--driver-memory`. En un paso, debería proporcionar los siguientes argumentos al paso:

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/spark/examples/
jars/spark-examples.jar 10
```

También puede ajustar la configuración que puede no tener un modificador integrado utilizando la opción `--conf`. Para obtener más información sobre otras opciones que se pueden ajustar, consulte el tema Dynamically loading Spark properties en la documentación de Apache Spark.

Visualización del historial de aplicaciones de Spark

Puede ver los detalles de la interfaz de usuario de Tez, las aplicaciones de YARN y Spark mediante la pestaña Interfaces de usuario de aplicaciones de la página de detalles de un clúster en la consola. Las interfaces de usuario (IU) de aplicaciones de Amazon EMR le facilitan la resolución de problemas y el análisis de los trabajos activos y el historial de trabajos.

Para obtener más información, consulte Ver el historial de aplicaciones en la Guía de administración de Amazon EMR.

Accede a la web de Spark UIs

Para ver la web UIs de Spark, siga los procedimientos para crear un túnel SSH o un proxy en la sección denominada Conectar al clúster de la Guía de administración de Amazon EMR y, a continuación, vaya al ResourceManager YARN de su clúster. Elija el enlace de la aplicación en Tracking UI (IU de seguimiento). Si la aplicación está en ejecución, verá. ApplicationMaster Esto

le lleva a la IU web del nodo principal de la aplicación en el puerto 20888 donde se encuentra el controlador. El controlador podría estar situado en el nodo principal del clúster si se ejecuta en modo de cliente de YARN. Si ejecuta una aplicación en el modo de clúster de YARN, el controlador se encuentra en el ApplicationMaster correspondiente a la aplicación del clúster. Si tu aplicación ha finalizado, verás el Historial, que te llevará al número de puerto de la HistoryServer interfaz de usuario de Spark, 18080, del nodo principal del clúster de EMR. Esto es para aplicaciones que ya se han completado. También puedes ir a la HistoryServer interfaz de usuario de Spark directamente en http: //:18080/. *master-public-dns-name*

Con la versión 5.25.0 y posteriores de Amazon EMR, puede acceder a la interfaz de usuario del servidor del historial de Spark desde la consola sin configurar un proxy web a través de una conexión SSH. Para obtener más información, consulte Ver interfaces de usuario de aplicaciones persistentes.

Uso del conector Amazon Kinesis Data Streams de streaming estructurado de Spark

Las versiones 7.1.0 y posteriores de Amazon EMR incluyen un conector de Amazon Kinesis Data Streams de streaming estructurado Spark en la imagen de la versión. Con este conector, puede usar Spark en Amazon EMR para procesar los datos que están almacenados en Amazon Kinesis Data Streams. El conector admite tanto los tipos de consumidores `GetRecords` (rendimiento compartido) como `SubscribeToShard` (distribución mejorada). Esta integración se basa en `spark-sql-kinesis-connector`. Para obtener información detallada sobre cómo empezar a utilizar el conector, consulte el README.

En el siguiente ejemplo, se muestra cómo utilizar el conector para iniciar una aplicación Spark con Amazon EMR.

```
spark-submit my_kinesis_streaming_script.py
```

Uso de la integración de Amazon Redshift para Apache Spark con Amazon EMR

Con la versión 6.4.0 y posteriores de Amazon EMR, la imagen de cada versión incluye un conector entre Apache Spark y Amazon Redshift. Con este conector, puede usar Spark en Amazon EMR para procesar los datos almacenados en Amazon Redshift. En el caso de las versiones 6.4.0 a 6.8.0 de

Amazon EMR, la integración se basa en el conector de código abierto `spark-redshift`. En el caso de las versiones 6.9.0 y posteriores de Amazon EMR, la integración de Amazon Redshift para Apache Spark se ha migrado de la versión de la comunidad a una integración nativa.

Temas

- Lanzamiento de una aplicación de Spark mediante la integración de Amazon Redshift para Apache Spark
- Autenticación con la integración de Amazon Redshift para Apache Spark
- Lectura y escritura desde y hacia Amazon Redshift
- Consideraciones y limitaciones al utilizar el conector de Spark

Lanzamiento de una aplicación de Spark mediante la integración de Amazon Redshift para Apache Spark

En el caso de las versiones 6.4 a 6.9 de Amazon EMR, debe usar la opción `--jars` o `--packages` para especificar cuáles de los siguientes archivos JAR desea usar. La opción `--jars` especifica las dependencias almacenadas localmente, en HDFS o mediante HTTP/S. Para ver otras ubicaciones de archivos compatibles con la opción `--jars`, consulte la sección Advanced Dependency Management en la documentación de Spark. La opción `--packages` especifica las dependencias almacenadas en el repositorio público de Maven.

- `spark-redshift.jar`
- `spark-avro.jar`
- `RedshiftJDBC.jar`
- `minimal-json.jar`

Las versiones 6.10.0 y posteriores de Amazon EMR no requieren la dependencia `minimal-json.jar` e instalan automáticamente las demás dependencias en cada clúster de forma predeterminada. En los siguientes ejemplos se muestra cómo lanzar una aplicación de Spark con la integración de Amazon Redshift para Apache Spark.

Amazon EMR 6.10.0 +

En el siguiente ejemplo se muestra cómo lanzar una aplicación de Spark con el conector `spark-redshift` con las versiones 6.10 y posteriores de Amazon EMR.

```
spark-submit my_script.py
```

Amazon EMR 6.4.0 - 6.9.x

Para lanzar una aplicación de Spark con el conector `spark-redshift` en las versiones 6.4 a 6.9 de Amazon EMR, debe usar la opción `--jars` o `--packages`, como se muestra en el siguiente ejemplo. Tenga en cuenta que las rutas enumeradas con la opción `--jars` son las rutas predeterminadas para los archivos JAR.

```
spark-submit \
  --jars /usr/share/aws/redshift/jdbc/RedshiftJDBC.jar,/usr/share/aws/redshift/
spark-redshift/lib/spark-redshift.jar,/usr/share/aws/redshift/spark-redshift/lib/
spark-avro.jar,/usr/share/aws/redshift/spark-redshift/lib/minimal-json.jar \
  my_script.py
```

Autenticación con la integración de Amazon Redshift para Apache Spark

AWS Secrets Manager Utilización para recuperar credenciales y conectarse a Amazon Redshift

El siguiente ejemplo de código muestra cómo recuperar credenciales AWS Secrets Manager para conectarse a un clúster de Amazon Redshift con la PySpark interfaz de Apache Spark en Python.

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

secretsmanager_client = boto3.client('secretsmanager')
secret_manager_response = secretsmanager_client.get_secret_value(
    SecretId='string',
    VersionId='string',
    VersionStage='string'
)
username = # get username from secret_manager_response
password = # get password from secret_manager_response
url = "jdbc:redshift://redshifthost:5439/database?user=" + username + "&password=" +
 password
```

```
# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3://path/for/temp/data") \
    .load()
```

Uso de IAM para recuperar credenciales y conectarse a Amazon Redshift

Puede utilizar el controlador JDBC versión 2 proporcionado por Amazon Redshift para conectarse a Amazon Redshift con el conector de Spark. Para usar AWS Identity and Access Management (IAM), configure la URL de JDBC para usar la autenticación de IAM. Para conectarse a un clúster de Redshift desde Amazon EMR, debe conceder a su rol de IAM permiso para recuperar credenciales de IAM temporales. Asigne los siguientes permisos a su rol de IAM para que pueda recuperar credenciales y ejecutar operaciones de Amazon S3.

- Redshift: GetClusterCredentials (para clústeres de Amazon Redshift aprovisionados)
- Redshift: DescribeClusters (para clústeres de Amazon Redshift aprovisionados)
- Redshift: GetWorkgroup (para grupos de trabajo sin servidor de Amazon Redshift)
- Redshift: GetCredentials (para grupos de trabajo sin servidor de Amazon Redshift)
- s3: GetBucket
- s3: GetBucketLocation
- s3: GetObject
- s3: PutObject
- s3: GetBucketLifecycleConfiguration

Para obtener más información sobre `GetClusterCredentials`, consulte Políticas de recursos de `GetClusterCredentials`.

También debe asegurarse de que Amazon Redshift puede asumir el rol de IAM durante las operaciones `COPY` y `UNLOAD`.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
```

```
            "Principal": {
                "Service": "redshift.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

En el siguiente ejemplo, se utiliza la autenticación de IAM entre Spark y Amazon Redshift:

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshift-host:redshift-port/db-name"
iam_role_arn = "arn:aws:iam::account-id:role/role-name"

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("aws_iam_role", iam_role_arn) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3a://path/for/temp/data") \
    .mode("error") \
    .load()
```

Lectura y escritura desde y hacia Amazon Redshift

Los siguientes ejemplos de código se utilizan PySpark para leer y escribir datos de muestra desde y hacia una base de datos de Amazon Redshift con una API de fuente de datos y mediante SparkSQL.

Data source API

Se utiliza PySpark para leer y escribir datos de muestra desde y hacia una base de datos de Amazon Redshift con una API de fuente de datos.

```
import boto3
from pyspark.sql import SQLContext
```

```
sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .load()

df.write \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName_copy") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .mode("error") \
    .save()
```

SparkSQL

Se utiliza PySpark para leer y escribir datos de muestra desde y hacia una base de datos de Amazon Redshift con SparkSQL.

```
import boto3
import json
import sys
import os
from pyspark.sql import SparkSession

spark = SparkSession \
    .builder \
    .enableHiveSupport() \
    .getOrCreate()

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

bucket = "s3://path/for/temp/data"
```

```
tableName = "tableName" # Redshift table name

s = f"""CREATE TABLE IF NOT EXISTS {tableName} (country string, data string)
    USING io.github.spark_redshift_community.spark.redshift
    OPTIONS (dbtable '{tableName}', tempdir '{bucket}', url '{url}', aws_iam_role
 '{aws_iam_role_arn}' ); """

spark.sql(s)

columns = ["country" ,"data"]
data = [("test-country","test-data")]
df = spark.sparkContext.parallelize(data).toDF(columns)

# Insert data into table
df.write.insertInto(tableName, overwrite=False)
df = spark.sql(f"SELECT * FROM {tableName}")
df.show()
```

Consideraciones y limitaciones al utilizar el conector de Spark

- Le recomendamos que active SSL para la conexión JDBC desde Spark en Amazon EMR a Amazon Redshift.
- Le recomendamos que administre las credenciales del clúster de Amazon Redshift en AWS Secrets Manager como práctica recomendada. Consulte Utilización AWS Secrets Manager para recuperar credenciales para conectarse a Amazon Redshift para ver un ejemplo.
- Le recomendamos que pase un rol de IAM con el parámetro `aws_iam_role` para el parámetro de autenticación de Amazon Redshift.
- El URI `tempdir` apunta a una ubicación de Amazon S3. Este directorio temporal no se limpia automáticamente y, por lo tanto, podría agregar costos adicionales.
- Tenga en cuenta las siguientes recomendaciones para Amazon Redshift:
  - Le recomendamos que bloquee el acceso público al clúster de Amazon Redshift.
  - Le recomendamos que active el registro de auditoría de Amazon Redshift.
  - Le recomendamos que active el cifrado en reposo de Amazon Redshift.
- Tenga en cuenta las siguientes recomendaciones para Amazon S3:
  - Le recomendamos que bloquee el acceso público a los buckets de Amazon S3.

- Le recomendamos que utilice el cifrado del servidor de Amazon S3 para cifrar los buckets de Amazon S3 utilizados.
- Le recomendamos que utilice las políticas de ciclo de vida de Amazon S3 para definir las reglas de retención del bucket de Amazon S3.
- Amazon EMR siempre verifica el código importado desde el código abierto a la imagen. Por motivos de seguridad, no admitimos los siguientes métodos de autenticación de Spark a Amazon S3:
  - Establecer las claves de AWS acceso en la clasificación de `hadoop-env` configuración
  - Codificación de las claves de AWS acceso en el `tempdir` URI

Para obtener más información sobre el uso del conector y sus parámetros compatibles, consulte los siguientes recursos:

- Integración de Amazon Redshift para Apache Spark en la Guía de administración de Amazon Redshift
- Repositorio comunitario de `spark-redshift` en GitHub

Historial de lanzamientos de Spark

En la siguiente tabla, se muestra la versión de Spark incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

> **Important**
>
> La versión 2.3.1 de Apache Spark, disponible a partir de la versión 5.16.0 de Amazon EMR, aborda CVE-2018-8024 y CVE-2018-1334. Se recomienda migrar las versiones anteriores de Spark a la versión 2.3.1 o posteriores de Spark.

Información de la versión de Spark

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-7.9.0	3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.8.0	3.5.4	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.7.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
		datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya rn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.6.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya rn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-7.5.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.4.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-7.3.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.2.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.36.2	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.1.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-7.0.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.15.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.14.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.13.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.12.0	3.4.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.11.1	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.11.0	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.10.1	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.10.0	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.9.1	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.9.0	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.8.1	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.8.0	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.7.0	3.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.36.1	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.36.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.6.0	3.2.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.35.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.5.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.4.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.3.1	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.3.0	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.2.1	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.2.0	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.1.1	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.1.0	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-6.0.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.0.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.34.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.33.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.33.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.32.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.32.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.31.1	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.31.0	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.2	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.30.1	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.0	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.29.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.28.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.28.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.27.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.27.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.26.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.25.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.24.1	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.24.0	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.23.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.23.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.22.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.21.2	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.21.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.21.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.20.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.20.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.19.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.19.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.18.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.18.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.2	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.17.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.16.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.16.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.15.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.15.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.14.2	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.14.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.14.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.13.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.13.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.12.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.12.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.4	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.11.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.11.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.10.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.10.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.9.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.9.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.8.3	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.2	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.8.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.7.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.7.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.6.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.6.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.5.4	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.3	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.5.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.4.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.4.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.3.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.3	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.2	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.2.1	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.0	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.1.1	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.1.0	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.3	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.2	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-5.0.1	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.0	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.6	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.9.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.9.2	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.1	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.8.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.8.1	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.0	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.4	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.7.3	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.7.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.5.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.4.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.3.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Etiqueta de versión de Amazon EMR	Versión de Spark	Componentes instalados con Spark
emr-4.2.0	1.5.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.1.0	1.5.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.0.0	1.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Apache Sqoop

> **Note**
>
> La última versión que incluyó Apache Sqoop fue la versión 7.4 de Amazon EMR. Apache Sqoop se eliminará de la versión 7.5 y versiones posteriores de Amazon EMR.

Apache Sqoop es una herramienta para la transferencia de datos entre Amazon S3, Hadoop, HDFS y las bases de datos RDBMS. Para obtener más información, consulte el sitio web de Apache Sqoop. Sqoop se incluye en la versión 5.0.0 y posteriores de Amazon EMR. Las versiones anteriores incluyen Sqoop como una aplicación de entorno aislado. Para obtener más información, consulte Versiones de lanzamiento de Amazon EMR 4.x.

Temas

- Información de la versión de Sqoop
- Consideraciones relacionadas con el uso de Sqoop en Amazon EMR
- Historial de versiones de Sqoop

Información de la versión de Sqoop

Versión de Sqoop para la versión 7.4.

En la siguiente tabla, se muestra la versión de Sqoop incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Sqoop.

Información sobre la versión de Spark para emr-7.4.0

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-7.4.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Versión de Sqoop para 6.15.0

En la tabla siguiente, se muestra la versión de Sqoop que se incluye en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Sqoop.

Para ver la versión de los componentes instalados con Sqoop en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Sqoop para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.15.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Versión de Sqoop para 5.36.2

En la tabla siguiente, se muestra la versión de Sqoop que se incluye en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Sqoop.

Para ver la versión de los componentes instalados con Sqoop en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Sqoop para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.36.2	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Consideraciones relacionadas con el uso de Sqoop en Amazon EMR

Tenga en cuenta los siguientes elementos al ejecutar Sqoop en Amazon EMR.

Uso de Sqoop con integración HCatalog

Sqoop en Amazon EMR admite la integración de Sqoop. HCatalog Cuando utilice Sqoop para escribir el resultado en una HCatalog tabla de Amazon S3, deshabilite la escritura directa de Amazon EMR estableciendo las `mapred.output.direct.NativeS3FileSystem` propiedades y en. `mapred.output.direct.EmrFileSystem` `false` Para obtener más información, consulte Usando HCatalog. Puede utilizar los comandos `-D mapred.output.direct.NativeS3FileSystem=false` y `-D mapred.output.direct.EmrFileSystem=false` de Hadoop. Si no se desactiva la escritura directa, no se produce ningún error, pero la tabla se crea en Amazon S3 y no se escriben datos.

Compatibilidad con JDBC y bases de datos de Sqoop

De forma predeterminada, Sqoop tiene instalado un controlador de MariaDB y PostgresSQL. El controlador de PostgresSQL instalado para Sqoop solo funciona en PostgreSQL 8.4. Para instalar un conjunto alternativo de conectores JDBC para Sqoop, conéctese al nodo principal del clúster e instálelos en `/usr/lib/sqoop/lib`. A continuación se indican los enlaces para diversos conectores JDBC:

- MariaDB: About MariaDB Connector/J.
- PostgreSQL: controlador de JDBC PostgreSQL.
- SQLServer: Descargue el controlador JDBC de Microsoft para SQL Server.
- MySQL: Download Connector/J
- Oracle: Obtener controladores de JDBC de Oracle y UCP del repositorio de Oracle Maven

Las bases de datos compatibles con Sqoop se muestran en la siguiente URL, http://sqoop.apache.org/docs/*version*/SqoopUserGuide.html #_supported_databases, donde *version* está la versión de Sqoop que está utilizando, por ejemplo, la 1.4.6. Si la cadena de conexión de JDBC no coincide con ninguna de las cadenas de esta lista, debe especificar un controlador.

Por ejemplo, puede exportar a una tabla de base de datos de Amazon Redshift con el siguiente comando (para JDBC 4.1):

```
sqoop export --connect jdbc:redshift://$MYREDSHIFTHOST:5439/mydb --
table mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 com.amazon.redshift.jdbc41.Driver --username master --password Mymasterpass1
```

Puede utilizar las cadenas de conexión de MariaDB y MySQL, pero si especifica la cadena de conexión de MariaDB, tiene que especificar el controlador:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Si utiliza el cifrado de Capa de conexión segura para acceder a la base de datos, tiene que utilizar un URI de JDBC como en el siguiente ejemplo de exportación de Sqoop:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb?
verifyServerCertificate=false&useSSL=true&requireSSL=true --table mysqoopexport --
```

```
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Para más información sobre el cifrado SSL en RDS, consulte Uso de SSL para cifrar una conexión a una instancia de base de datos en la Guía del usuario de Amazon RDS.

Para obtener más información, consulte la documentación de Apache Sqoop.

Cómo proteger la contraseña

Hay varios métodos entre los cuales podría elegir para asegurarse de transferir su contraseña de forma segura:

Java KeyStore

El método preferido cifra la contraseña con un código Java KeyStore (JKS), lo que elimina la necesidad de almacenar la contraseña en un formato legible.

1. Cree un alias de contraseña. Cuando aparezca el mensaje, ingrese su contraseña para acceder a la base de datos.

   ```
   hadoop credential create mydb.password.alias -provider jceks://hdfs/user/root/
   mysql.password.jceks
   ```

2. Utilice el alias de contraseña para iniciar el trabajo de Sqoop:

   ```
   sqoop export -Dhadoop.security.credential.provider.path=jceks://hdfs/user/
   root/mysql.password.jceks --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
    mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
    org.mariadb.jdbc.Driver --username master --password-alias mydb.password.alias
   ```

--password-file

También puede usar el comando `--password-file` para transferir la contraseña a través de un archivo, como se muestra en el siguiente ejemplo:

1. Cree un nuevo archivo que contenga la contraseña:

   ```
   echo -n 'Mymasterpass1' > /home/hadoop/mysql-pass.password
   ```

2. Utilice el archivo para iniciar el trabajo de Sqoop:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
 mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 org.mariadb.jdbc.Driver --username master --password-file /home/hadoop/mysql-
pass.password
```

-P

También puede usar el comando `-P` para escribir la contraseña, como se muestra en el siguiente ejemplo:

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver
 --username master -P
```

Historial de versiones de Sqoop

En la siguiente tabla, se muestra la versión de Sqoop incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Sqoop

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-7.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
		timeline-server, mariadb-server, sqoop-client
emr-7.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.36.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-7.0.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.13.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.12.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.11.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.11.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.10.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.10.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.9.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.9.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.8.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.8.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.7.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.36.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.36.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.6.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.35.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.5.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.3.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.2.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-6.1.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.34.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.33.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.33.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.32.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.32.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.31.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.31.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.30.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.29.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.28.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.28.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.27.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.27.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.26.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.25.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.24.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.24.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.23.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.23.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.22.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.21.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.21.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.21.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.20.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.20.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.17.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.16.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.16.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.15.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.14.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.14.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.13.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.13.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.12.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.12.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.11.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.11.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.10.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.10.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.9.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.5.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.5.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.3.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.2.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.2.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.0.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Etiqueta de versión de Amazon EMR	Versión de Sqoop	Componentes instalados con Sqoop
emr-5.0.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

TensorFlow

TensorFlow es una biblioteca matemática simbólica de código abierto para aplicaciones de inteligencia artificial y aprendizaje profundo. Para obtener más información, consulte el TensorFlow sitio web. TensorFlow está disponible con la versión 5.17.0 y posteriores de Amazon EMR.

En la tabla siguiente, se muestra la versión TensorFlow incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes con los que Amazon EMR se instala. TensorFlow

Para ver la versión de los componentes instalados con TensorFlow esta versión, consulte Versiones de los componentes de la versión 7.9.0.

TensorFlow información de la versión para emr-7.9.0

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-7.9.0	TensorFlow 2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

En la tabla siguiente, se muestra la versión TensorFlow incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes con los que Amazon EMR se instala. TensorFlow

Para ver la versión de los componentes instalados con TensorFlow esta versión, consulte Versiones de los componentes de la versión 6.15.0.

TensorFlow información de la versión para emr-6.15.0

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.15.0	TensorFlow 2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

En la tabla siguiente, se muestra la versión TensorFlow incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes con los que Amazon EMR se instala. TensorFlow

Para ver la versión de los componentes instalados con TensorFlow esta versión, consulte Versiones de los componentes de la versión 5.36.2.

TensorFlow información de la versión para emr-5.36.2

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.36.2	TensorFlow 2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

TensorFlow compilaciones por tipo de EC2 instancia de Amazon

Amazon EMR utiliza diferentes compilaciones de la TensorFlow biblioteca en función de los tipos de instancias que elija para el clúster. Amazon EMR admite clústeres con tipos de instancia aarch64 (Graviton) TensorFlow para EMR-7.5.0 y versiones posteriores. En la siguiente tabla, se muestran las compilaciones por tipo de instancia.

EC2 tipos de instancias	TensorFlow construir
M5 y C5	Tensorflow 2.16.1 con optimización MKL de Intel
P2, P4D, P5, G4DN, G5, G6 y GR6	Tensorflow 2.16.1 con CUDA 12.3, cuDNN 8.9.7.29
P3, P3DN, G3 y G3S	Tensorflow 2.16.1 con CUDA 12.3, cuDNN 8.9.7.29, NCCL 2.20.3-1 Nvidia NCCL solo está disponible en las instancias P3. Contrato de licencia para el usuario final (CLUF): al utilizar componentes Nvidia en Amazon EMR, acepta los términos y condiciones que se detallan en el CLUF del producto.
Todas las demás, excepto las instancias de Graviton	Tensorflow 2.16.1

Seguridad

Además de seguir las instrucciones que figuran en Uso TensorFlow seguro, recomendamos lanzar el clúster en una subred privada para ayudarle a limitar el acceso a los orígenes de confianza. Para más información, consulte Opciones de Amazon VPC en la Guía de administración de Amazon EMR.

Usando TensorBoard

TensorBoard es un conjunto de herramientas de visualización para TensorFlow programas. Para más información, consulte TensorBoard: aprendizaje visualizado en el sitio web de TensorFlow.

Para utilizarlo TensorBoard con Amazon EMR, debe empezar TensorBoard en el nodo maestro del clúster.

Para utilizar TensorBoard con TensorFlow en Amazon EMR

1. Conecte al nodo principal del clúster utilizando SSH. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.
2. Escriba el comando siguiente para iniciar Tensorboard en el nodo principal. Sustituya */my/log/directory* por un directorio del nodo principal donde haya generado y almacenado datos de resumen utilizando un generador de resúmenes.

   Amazon EMR 5.19.0 and later

   ```
   python3 -m tensorboard.main --logdir=/home/hadoop/tensor --bind_all
   ```

   Amazon EMR 5.18.1 and earlier

   ```
   python3 -m tensorboard.main --logdir=/my/log/dir
   ```

   De forma predeterminada, el nodo maestro aloja TensorBoard mediante el puerto 6006 y el nombre de DNS público maestro. Tras empezar TensorBoard, el resultado de la línea de comandos presenta la URL a la que se puede utilizar para conectarse TensorBoard, como se muestra en el siguiente ejemplo:

   ```
   TensorBoard 2.16.1 at http://master-public-dns-name:6006 (Press CTRL+C to quit)
   ```
3. Configure el acceso a las interfaces web en el nodo principal desde los clientes de confianza. Para más información, consulte Ver interfaces web alojadas en los clústeres de Amazon EMR en la Guía de administración de Amazon EMR.
4. Abrir TensorBoard en`http://master-public-dns-name:6006`.

TensorFlow historial de lanzamientos

En la tabla siguiente, se muestra la versión TensorFlow incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

TensorFlow información sobre la versión

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-7.9.0	2,1161	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.8.0	2,1161	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.7.0	2,1161	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.6.0	2,1161	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.5.0	2,1161	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.4.0	2,1161	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-7.3.0	2,1161	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.2.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.2	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-7.1.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.0.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.15.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.14.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.13.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.12.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.11.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.11.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.10.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.10.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.1	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.0	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.8.1	2,9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.8.0	2,9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.7.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.36.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.6.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.35.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.5.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.4.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.3.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.3.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.2.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.2.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-6.0.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.0.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.34.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.33.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.33.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.32.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.32.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.30.2	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.29.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.27.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.27.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.26.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.25.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.23.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.23.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.22.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.21.2	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.20.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.20.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.19.1	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.19.0	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Etiqueta de versión de Amazon EMR	TensorFlow Versión	Componentes instalados con TensorFlow
emr-5.17.2	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Apache Tez

Apache Tez es un marco que crea un gráfico acíclico dirigido (DAG) complejo de tareas para el procesamiento de datos. Puede usarlo como una alternativa a Hadoop MapReduce para algunos casos de uso. Por ejemplo, puede ejecutar flujos de trabajo de Pig and Hive con Hadoop MapReduce o puede usar Tez como motor de ejecución. Para obtener más información, consulte https://tez.apache.org/. La versión 4.7.0 y posteriores de Amazon EMR incluyen Tez.

En la siguiente tabla, se muestra la versión de Tez incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Tez.

Para ver la versión de los componentes instalados con Tez en esta versión, consulte Versiones de componentes de la versión 7.9.0.

Información sobre la versión de Tez para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-7.9.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

En la tabla siguiente, se muestra la versión de Tez incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Tez.

Para ver la versión de los componentes instalados con Tez en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Tez para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.15.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

En la tabla siguiente, se muestra la versión de Tez incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Tez.

Para ver la versión de los componentes instalados con Tez en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Tez para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.36.2	Tez 0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Temas

- Creación de un clúster con Tez
- Configuración de Tez
- IU web de Tez
- Timeline Server
- Historial de versiones de Tez

Creación de un clúster con Tez

Para instalar Tez, elija Apache Tez como aplicación al crear el clúster.

Para crear un clúster con Tez instalado utilizando la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Elija Crear clúster e Ir a las opciones avanzadas.
3. Debajo de Configuración de software, seleccione una Versión de emr-4.7.0 o superior.
4. Seleccione Tez junto con otras aplicaciones que desee que Amazon EMR instale.
5. Seleccione las demás opciones que necesite y, a continuación, elija Create cluster (Crear clúster).

Para crear un clúster con Tez mediante el AWS CLI

- Utilice el comando `create-cluster` junto con la opción `-- applications` para especificar Tez. El siguiente ejemplo crea un clúster con Tez instalado.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Cluster with Tez" --release-label emr-7.9.0 \
```

```
--applications Name=Tez --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Configuración de Tez

Puede personalizar Tez configurando valores que utilicen la clasificación de configuración `tez-site` cuando cree el clúster, que configura los ajustes del archivo de configuración `tez-site.xml`. Para obtener más información, consulte TezConfigurationla documentación de Apache Tez. Para cambiar Hive o Pig para utilizar el motor de ejecución de Tez, utilice las clasificaciones de configuración de `hive-site` y `pig-properties`, según proceda. A continuación se muestran ejemplos.

Configuración de ejemplo

Example Ejemplo: Personalizar el nivel de registro raíz de Tez y configurar Tez como el motor de ejecución para Hive y Pig

El comando `create-cluster` de ejemplo que se muestra a continuación crea un clúster con Tez, Hive y Pig instalados. El comando hace referencia a un archivo almacenado en Amazon S3, `myConfig.json`, que especifica las propiedades de la clasificación `tez-site` que establece `tez.am.log.level` en `DEBUG`, y establece el motor de ejecución en Tez para Hive y Pig mediante las clasificaciones de configuración `hive-site` y `pig-properties`.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Tez Name=Hive Name=Pig --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json --
use-default-roles
```

A continuación, se muestra contenido de ejemplo de `myConfig.json`.

```
[
```

```
    {
      "Classification": "tez-site",
      "Properties": {
        "tez.am.log.level": "DEBUG"
      }
    },
    {
      "Classification": "hive-site",
      "Properties": {
        "hive.execution.engine": "tez"
      }
    },
    {
      "Classification": "pig-properties",
      "Properties": {
        "exectype": "tez"
      }
    }
  ]
```

> **Note**
>
> Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

Apertura de división asíncrona de Tez

Cuando hay una gran cantidad de archivos pequeños en la ruta de la tabla y una consulta intenta leerlos todos, cada archivo pequeño que corresponde a cada división individual se combina en una división agrupada de Tez. A continuación, un solo asignador procesa la única división agrupada de Tez. Dado que la ejecución es sincrónica, cada división individual dentro de la división agrupada se procesa una por una. Esto requiere que `RecordReader` los objetos deben procesar las divisiones de manera sincronizada.

Amazon EMR 6.15.0 presenta configuraciones que puede especificar para abrir de forma asíncrona las divisiones de entrada en una división agrupada de Tez. TEZ-4397 inició la característica, pero se

produjeron regresiones en OSS Hive. EMR Hive corrigió las regresiones y errores adicionales en la tabla ACID de Hive. Esta mejora se traduce en un rendimiento más rápido de las consultas de lectura cuando hay una gran cantidad de divisiones de entrada en una sola división agrupada de Tez.

Nombre	Clasificación	Descripción
`tez.grouping.split.init.threads`	`tez-site`	Especifica la cantidad de subprocesos daemon que utiliza Tez para realizar el inicio previo de `RecordReaders` y la apertura de las divisiones. En el caso de las tablas ACID, el valor máximo admitido de `tez.grouping.split.init.threads` es 1.
`tez.grouping.split.init.recordreaders`	`tez-site`	Especifica la cantidad de `RecordReaders` subprocesos daemon que deben mantener previamente iniciados. Esto puede ser de ayuda cuando la división agrupada de Tez contiene una gran cantidad de `InputSplits`. La inicialización de `RecordReaders` para procesar esas divisiones de entrada se puede realizar de forma asíncrona con subprocesos daemon en lugar de recurrir a un procesamiento secuencial.

Notas de configuración:

Asunto	Detalles
Parámetros de configuración recomendados	Se recomienda establecer los valores de configuración anteriores en los valores que desee tanto en como `hive-site` `entez-site`.
Valores coincidentes	Los valores de los ajustes de configuración deben ser los mismos en `hive-site` y `tez-site`.
Recomendación del LLAP	No se recomienda utilizar esta función cuando el LLAP está activado.

Evaluación comparativa de la apertura de división asíncrona de Tez

Utilizamos los siguientes entornos y configuraciones para realizar una evaluación comparativa de la capacidad de apertura de división asíncrona de Tez:

- Entorno de referencia: clúster de Amazon EMR con 1 nodo principal que utiliza m5.16xlarge y 16 nodos centrales que utilizan m5.16xlarge.
- Configuraciones de evaluación comparativa: para simular el escenario de evaluación comparativa donde hay una gran cantidad de divisiones de entrada en una sola división agrupada de Tez, `tez.grouping.split-count` se establece en `1`.
- Tabla utilizada para la evaluación comparativa: la tabla contiene 200 particiones, y cada partición contiene un único archivo. El punto de referencia se realiza cuando esa tabla contiene archivos CSV y cuando esa tabla contiene archivos de parquet. Consulta de Hive para la evaluación comparativa: `SELECT COUNT(*)` a partir de la tabla diez veces y tome el tiempo de ejecución promedio.
- Configuraciones para habilitar la apertura de división asíncrona de Tez: como se indica a continuación:
  - `tez.grouping.split.init.threads` = `4`
  - `tez.grouping.split.init.recordreaders` = `10`

Conjunto de datos	Característica deshabilitada (base de referencia)	Característica habilitada	Mejora
Conjunto de datos CSV	90,26 segundos	79,20 segundos	12,25%
Conjunto de datos Parquet	54,67 segundos	42,23 segundos	22,75%

IU web de Tez

Tez tiene su propia interfaz de usuario web. Para ver la interfaz de usuario web, consulte la siguiente URL.

```
http://masterDNS:8080/tez-ui
```

Para habilitar la pestaña Hive Queries en la interfaz de usuario web de Tez, establezca la siguiente configuración.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.exec.pre.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.post.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.failure.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook"
    }
  }
]
```

También puede ver los detalles de la interfaz de usuario de las aplicaciones Tez, Spark y YARN mediante los enlaces de la pestaña Interfaces de usuario de aplicaciones de la página de detalles de un clúster en la consola. Las interfaces de usuario (UI) de aplicaciones de Amazon EMR se alojan fuera del clúster y están disponibles una vez terminado el clúster. No requieren que configure una conexión SSH o un proxy web, lo que facilita la solución de problemas y el análisis de trabajos activos y el historial de trabajos.

Para obtener más información, consulte Ver el historial de aplicaciones en la Guía de administración de Amazon EMR.

Timeline Server

El servidor YARN Timeline se ha configurado para ejecutarse cuando se instala Tez. Para ver los trabajos enviados a través de Tez o de los motores de MapReduce ejecución que utilizan el Timeline Server, consulte la interfaz de usuario web mediante la URL`http://master-public-DNS:8188`. Para más información, consulte Ver interfaces web alojadas en los clústeres de Amazon EMR en la Guía de administración de Amazon EMR.

Historial de versiones de Tez

En la siguiente tabla, se muestra la versión de Tez incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Tez

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-7.9.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.8.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
		hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.7.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.6.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-7.5.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.4.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.3.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-7.2.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-5.36.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-7.1.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-7.0.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.15.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.14.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.13.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.12.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.11.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.11.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.9.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.9.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.8.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.8.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.7.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.36.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.36.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.6.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.35.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.5.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.4.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.3.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.3.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.1.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.1.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.0.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-6.0.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.34.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.33.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.33.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.31.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.31.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.30.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.29.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.28.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.28.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.27.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.27.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.26.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.25.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.24.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.24.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.23.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.23.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.22.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.2	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.21.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.20.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.20.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.18.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.18.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.17.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.16.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.16.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.14.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.13.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.13.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.12.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.11.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.11.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.10.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.10.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.6.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.6.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.5.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.5.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.3.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.2.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.2.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.0.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-5.0.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.6	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-4.9.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-4.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-4.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-4.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.4	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.3	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Etiqueta de versión de Amazon EMR	Versión de Tez	Componentes instalados con Tez
emr-4.7.2	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.1	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.0	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Notas de la versión de Tez por versión

Temas

- Amazon EMR 7.8.0: notas de la versión de Tez
- Amazon EMR 7.7.0: notas de la versión de Tez
- Amazon EMR 7.6.0 - Notas de la versión de Tez
- Amazon EMR 7.5.0: notas de la versión de Tez
- Amazon EMR 7.4.0 - Notas de la versión de Tez
- Amazon EMR 7.2.0 - Notas de la versión de Tez
- Amazon EMR 7.1.0: notas de la versión de Tez
- Amazon EMR 7.0.0 - Notas de la versión de Tez
- Amazon EMR 6.15.0: notas de la versión de Tez
- Amazon EMR 6.14.0: notas de la versión de Tez
- Amazon EMR 6.13.0: notas de la versión de Tez
- Amazon EMR 6.12.0: notas de la versión de Tez
- Amazon EMR 6.11.0: notas de la versión de Tez
- Amazon EMR 6.10.0: notas de la versión de Tez
- Amazon EMR 6.9.0: notas de la versión de Tez
- Amazon EMR 6.8.0: notas de la versión de Tez
- Amazon EMR 6.7.0: notas de la versión de Tez
- Amazon EMR 6.6.0: notas de la versión de Tez

Amazon EMR 7.8.0: notas de la versión de Tez

Amazon EMR 7.8.0: cambios en el té

Tipo	Descripción
Mejoras	Se actualizó la versión bouncycastle a 1.78 para que coincida con Hadoop

Amazon EMR 7.7.0: notas de la versión de Tez

Amazon EMR 7.7.0: cambios en el té

Tipo	Descripción
Mejoras	La `tez.task.relaxed.locality` propiedad de Apache Tez controla si la programación de tareas sigue estrictamente las restricciones de localidad de los datos (localida d del rack y del nodo). Cuando se establece en `true` (predeterminado en EMR-7.6+), Tez no exige la localidad, lo que permite asignar las tareas a cualquier contenedor disponible, lo que mejora la utilización de los recursos y reduce los tiempos de espera en los clústeres ocupados.

Amazon EMR 7.6.0 - Notas de la versión de Tez

Amazon EMR 7.6.0: cambios en el té

Tipo	Descripción
Mejoras	Si se cambia el valor predeterminado de false, no se tendrán en cuenta las restricciones de localidad del rack y el nodo al solicitar un contenedor para una tarea `tez.task. relaxed.locality`
Mejora	Ajuste las configuraciones para deshabilitar el retraso debido a la localidad y permitir el respaldo no local
Mejora	TEZ-4547: Añada Tez AM JoBid al JobConf

Amazon EMR 7.5.0: notas de la versión de Tez

Amazon EMR 7.5.0: cambios en el té

Tipo	Descripción
Mejoras	Se habilitó la preferencia sobre la tarea de forma predeterminada
Mejora	Se aumentó a 10 el número de lectores de registros que se inician de forma asíncrona y proactiva
Mejora	TEZ-4580: Cuando se habilita la reutilización, se evita la aparición de nuevos contenedores

Amazon EMR 7.4.0 - Notas de la versión de Tez

Amazon EMR 7.4.0: cambios en el té

Tipo	Descripción
Mejoras	Eliminar el EventCounter apéndice de Tez eliminado en Hadoop 3.4.0 mediante HADOOP-17524
Upgrade	Actualice commons-cli a la versión 1.5.0
Upgrade	Actualiza la versión bouncy-castle a la 1.70.0

Amazon EMR 7.2.0 - Notas de la versión de Tez

Amazon EMR 7.2.0: cambios en el té

Tipo	Descripción
Mejoras	Se proporciona una opción para deshabilitar las restricciones de localidad al solicitar contenedo res para tareas.
Corrección de errores	Haz que TaskDependencies# sea seguro para addTaskDependency subprocesos
Upgrade	Actualice la versión TLS a 1.3 para Tez.

Amazon EMR 7.2.0: nuevas configuraciones

Clasificación	Nombre	Predeterminado/a	Descripción
tez-site	tez.task.relaxed.l ocality	false	Cuando está habilitad a, no se tienen en cuenta las restricci ones de localidad de los racks y los nodos al solicitar un contenedor para una tarea.

Amazon EMR 7.1.0: notas de la versión de Tez

Amazon EMR 7.1.0: cambios en Tez

Tipo	Descripción
Corrección de errores	TEZ-4394 — Netty4 ShuffleHandler: debería usar 1 hilo principal (#256)

Tipo	Descripción
Upgrade	Actualiza Netty a la versión 4.1.100.Final.
Upgrade	Actualización de Jetty a 9.4.53.v20231009.

Amazon EMR 7.0.0 - Notas de la versión de Tez

Amazon EMR 7.0.0 Tez contiene todos los cambios y actualizaciones realizados hasta Amazon EMR-6.15.0 Tez, inclusive.

Amazon EMR 7.0.0: cambios en el té

Tipo	Descripción
Upgrade	Tez Runtime ahora usa Java 17 de forma predeterminada. Consulte la Guía de versión de EMR 7.0.0 para obtener más información.

Amazon EMR 6.15.0: notas de la versión de Tez

Amazon EMR 6.15.0: cambios en Tez

Tipo	Descripción
Característica	TEZ-4397: apertura de las divisiones de entrada de Tez de manera asíncrona
Upgrade	TEZ-4493: actualización de Apache Hadoop a la versión 3.3.6

Amazon EMR 6.15.0: características de Tez

- Apertura de división asíncrona de Tez: Amazon EMR 6.15.0 presenta configuraciones que puede especificar para abrir de forma asíncrona las divisiones de entrada en una división agrupada de Tez. TEZ-4397 inició la característica, pero se produjeron regresiones en OSS Hive. Amazon

EMR Hive corrigió las regresiones y errores adicionales en la tabla ACID de Hive. Esta mejora resulta en un rendimiento más rápido de las consultas de lectura cuando hay un gran número de divisiones de entrada en una única división agrupada de Tez. Para obtener más información, consulte Apertura de división asíncrona de Tez.

Amazon EMR 6.14.0: notas de la versión de Tez

Amazon EMR 6.14.0: cambios en Tez

Tipo	Descripción
Mejoras	Actualización de la versión de TLS en Tez a 1.2

Amazon EMR 6.13.0: notas de la versión de Tez

Amazon EMR 6.13.0: cambios en Tez

Tipo	Descripción
Corrección de errores	REVERT TEZ-4295: no se pudieron descompri mir los datos. La longitud del búfer es demasiado pequeña.
Corrección de errores	REVERT TEZ-4302: `NullPointerExcepti on en CodecUtils con GzipCodec .`
Corrección de errores	REVERT TEZ-4234: el compresor puede provocar `IllegalArgumentException` en `Buffer.limit` donde el límite supera la capacidad.
Corrección de errores	REVERT TEZ-4135: se mejora la asignación de memoria al ejecutar lecturas en memoria.

Amazon EMR 6.12.0: notas de la versión de Tez

Amazon EMR 6.12.0: cambios en Tez

Tipo	Descripción
Mejoras	Se ha agregado compatibilidad con JDK 11 y JDK 17 Runtime
Corrección de errores	TEZ-4492: se actualiza Bowerrc para utilizar el espejo bower.herokuapp para evitar el problema CERT_EXPIRE del registro de Bower (BOWER-2608)
Upgrade	Se ha actualizado Surefire a la versión 3.0.0-M7

Amazon EMR 6.11.0: notas de la versión de Tez

Amazon EMR 6.11.0: cambios en Tez

Tipo	Descripción
Error	Se ha corregido la transición de estado de vértice no válida durante la limpieza en el nivel de vértice de los datos de mezclas aleatorias
Error	Se ha corregido el problema por el que la limpieza en el nivel de vértice o DAG de los datos de mezclas aleatorias no funcionaba
Mejora	Activa tez.am.dag.cleanup.on.completion de forma predeterminada para borrar los datos de reproducción aleatoria si se ha completado DAGs

Amazon EMR 6.10.0: notas de la versión de Tez

Amazon EMR 6.10.0: cambios en Tez

Tipo	Descripción
Característica	`tez.runtime.transfer.data-via-events.enabled` habilitado de forma predeterminada
Portabilidad con versiones anteriores	TEZ-4450: se corrige un error al recuperar datos de mezclas aleatorias cuando estos datos se transfieren mediante eventos de movimiento de datos
Portabilidad con versiones anteriores	TEZ-4460: se corrige un error de tiempo de espera de lectura al obtener datos aleatorios desde el Tez Shuffle Handler
Portabilidad con versiones anteriores	TEZ-4455: Se ha añadido una canalización ShuffleHandler para mejorar la LoggingHandler capacidad de depuración
Error	Se ha corregido el problema por el que la tarea de Tez se bloqueaba de forma intermitente cuando la prioridad de la tarea estaba habilitad a

Amazon EMR 6.9.0: notas de la versión de Tez

Amazon EMR 6.9.0: cambios en Tez

Tipo	Descripción
Upgrade	Se ha actualizado Tez a la versión 0.10.2. Para más información, consulte el registro de cambios de Apache Tez 0.10.2.

Tipo	Descripción
Upgrade	Se ha actualizado Hadoop a la versión 3.3.3.
Error	Se ha deshabilitado `tez.runtime.transfer.data-via-events.enabled` de forma predeterminada debido a TEZ-4450.

Amazon EMR 6.8.0: notas de la versión de Tez

Amazon EMR 6.8.0: cambios en Tez

Tipo	Descripción
Portabilidad con versiones anteriores	TEZ-3363: se eliminan los datos intermedios en el nivel de vértice para el controlador de mezclas aleatorias
Portabilidad con versiones anteriores	TEZ-4129: se eliminan los datos de intentos intermedios de los intentos fallidos del controlador de mezclas aleatorias
Portabilidad con versiones anteriores	TEZ-4430: se ha corregido la propiedad tez.task.launch.cmd-opts que no funcionaba

Amazon EMR 6.7.0: notas de la versión de Tez

Amazon EMR 6.7.0: cambios en Tez

Tipo	Descripción
Portabilidad con versiones anteriores	TEZ-4403: Actualice la versión J a la 1.7.36 SLF4
Portabilidad con versiones anteriores	TEZ-4405: se sustituye log4j 1.x por reload4j

Tipo	Descripción
Portabilidad con versiones anteriores	TEZ-4411: Error de compilación de Tez: no se encontró el archivo.js FileSaver

Amazon EMR 6.6.0: notas de la versión de Tez

Amazon EMR 6.6.0: cambios en Tez

Tipo	Descripción
Portabilidad con versiones anteriores	TEZ-3918: se ha corregido la propiedad tez.task.log.level que no funcionaba.
Portabilidad con versiones anteriores	TEZ-4353: se actualiza commons-io a la versión 2.8.0.
Portabilidad con versiones anteriores	TEZ-4114: se elimina la dependencia directa de Jetty de Tez.
Portabilidad con versiones anteriores	TEZ-4323: los archivos JAR de Jetty se han eliminado del paquete de distribución con TEZ-4114.

Trino

Trino es un motor de consultas de código abierto diseñado para consultas interactivas en una amplia gama de fuentes de datos. Estos pueden incluir bases de datos relacionales, datos basados en archivos, datos HDFS y otros. El propósito más común de Trino con Amazon EMR es ejecutar consultas SQL complejas en conjuntos de datos de gran tamaño almacenados en Amazon S3. También es compatible con ANSI SQL, lo que hace que sea familiar para los ingenieros de bases de datos, los analistas de datos y los científicos de datos que estén familiarizados con SQL.

> **Note**
>
> PrestoSQL pasó a llamarse Trino en diciembre de 2020. Las versiones 6.4.0 y posteriores de Amazon EMR generalmente hacen referencia a Trino, mientras que las versiones anteriores hacen referencia a PrestoSQL.

> **Important**
>
> PrestosQL, la versión anterior de Trino, todavía está disponible para su uso con Amazon EMR. Sin embargo, recomendamos encarecidamente utilizar Trino en el futuro con Amazon EMR. Tenga en cuenta también que Trino y PrestosQL no se pueden ejecutar simultáneamente en el mismo clúster.

En la siguiente tabla, se muestra la versión de Trino incluida en la última versión de Amazon EMR, junto con los componentes que Amazon EMR instala con Trino. Para ver la versión de los componentes instalados con Trino en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Trino para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Trino	Componentes instalados con Trino
emr-7.9.0	Trino 467	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Etiqueta de versión de Amazon EMR	Versión de Trino	Componentes instalados con Trino
		hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

Temas

- Historia y diseño de Trino
- Introducción a Trino
- Configuración de Trino en Amazon EMR
- Prácticas recomendadas para Trino en Amazon EMR
- Consideraciones sobre Trino
- Historial de lanzamientos de Trino

Historia y diseño de Trino

Trino está especializado en consultar grandes conjuntos de datos de muchas fuentes diferentes. Trino puede acceder a HDFS y consultarlo en un caso de uso tradicional de macrodatos, pero también puede consultar fuentes adicionales, como bases de datos relacionales y bases de datos NoSQL. Trino comenzó originalmente como una bifurcación del motor de consultas Presto, en 2019. Desde entonces, se ha desarrollado de forma independiente del código base de Presto.

Para obtener más información sobre el motor de consultas de Trino y cómo se utiliza, consulte el sitio web de Trino. Para leer la documentación original de Trino, consulte la descripción general de Trino.

Conceptos arquitectónicos

Trino puede ejecutar consultas rápidas y eficientes porque procesa los datos en paralelo en un clúster. Se ha diseñado teniendo en cuenta las consultas de un lago de datos, ya que está

especializado para consultas sobre grandes volúmenes de datos, normalmente en casos de uso relacionados con Hadoop y HDFS. Pero también puede consultar bases de datos relacionales tradicionales. Para obtener más información, consulte Arquitectura en la documentación de Trino.

Componentes de Trino

Trino tiene algunos componentes de arquitectura clave que funcionan juntos para que las consultas se ejecuten rápidamente. Es útil tener un conocimiento práctico de estos elementos a la hora de ajustar el clúster para obtener un mejor rendimiento:

- El coordinador es responsable de la organización de las consultas. Analiza y optimiza las consultas SQL entrantes, genera planes de ejecución, asigna tareas a los nodos de trabajo y recopila y reúne los resultados de las consultas. Además, supervisa el uso de los recursos y rastrea el estado de los nodos de trabajo. Para obtener más información, consulte Coordinator en la documentación de Trino.
- Los nodos de trabajo se encargan del procesamiento de datos para las consultas. Una vez que el coordinador asigna las tareas, los trabajadores recuperan los datos, realizan las operaciones necesarias, como uniones y agregaciones, e intercambian datos intermedios con otros trabajadores. Para obtener más información, consulte Worker en la documentación de Trino.
- Los conectores son complementos que permiten a Trino conectarse con varias fuentes de datos y consultarlas. Cada conector sabe cómo acceder a los datos de su origen y recuperarlos, como Amazon S3, Apache Hive o bases de datos relacionales. Estos conectores asignan los datos de origen a la estructura del esquema de Trino.
- Un catálogo es una colección lógica de esquemas y tablas asociados a un conector específico. Definidos en el coordinador, los catálogos permiten a Trino tratar diferentes fuentes de datos como un único espacio de nombres. Esto permite a los usuarios consultar varias fuentes juntas, como Hive y MySQL, de forma unificada en la misma consulta.
- Los clientes, como la CLI de Trino, se conectan mediante controladores JDBC y ODBC al coordinador de Trino para enviar consultas SQL. El coordinador gestiona el ciclo de vida de las consultas y proporciona los resultados al cliente para su posterior análisis o elaboración de informes.

Ejecución de consultas

Para entender cómo Trino toma las sentencias SQL y las ejecuta como consultas, consulte los conceptos de Trino en la documentación de Trino.

Introducción a Trino

Los procedimientos de esta sección muestran cómo configurar un clúster de Amazon EMR para consultar fuentes de datos de metastore con Trino. Estos metaalmacenes, que incluyen el catálogo de datos de AWS Glue, almacenan metadatos y objetos de bases de datos y administran los permisos de acceso. Los procedimientos incluyen los requisitos previos, los valores de configuración recomendados, la creación de conectores y la ejecución de consultas en las tablas de metastore.

Temas

- Complete los pasos previos para usar Amazon EMR con Trino
- Lanzar un clúster de Amazon EMR con Trino
- Conéctese al nodo principal del clúster de Amazon EMR y ejecute consultas

Complete los pasos previos para usar Amazon EMR con Trino

Si no ha utilizado AWS o no ha creado un clúster de Amazon EMR, complete estos pasos previos antes de crear un clúster de Amazon EMR con Trino.

AWS configuración del entorno

Complete estos pasos para configurar su AWS cuenta si aún no lo ha hecho:

1. Cree una AWS cuenta, si aún no tiene ninguna. Para obtener más información, consulta Crear una AWS cuenta en la Guía de referencia de administración de AWS cuentas.
2. Inicie sesión en su cuenta de como usuario administrativo.
3. Crea un grupo y asígnale usuarios.
4. Crea un par de EC2 claves de Amazon, que puedas usar más adelante para proteger la comunicación entre los recursos con SSH. Este paso es obligatorio si tiene previsto conectarse al nodo principal para llevar a cabo tareas. Para obtener más información, consulte Conectarse al nodo principal del clúster de Amazon EMR mediante SSH.

Lanzar un clúster de Amazon EMR con Trino

A continuación, se muestran las opciones de configuración correctas al crear un clúster con Trino.

Uso de un conector Hive para que los datos estén disponibles para su consulta

Puede configurar un conector Trino para un metaalmacén de Hive con el fin de consultar los datos del metaalmacén de su clúster. Un metaalmacén es una capa de abstracción que hace que el contenido o los datos basados en archivos estén disponibles en forma de tablas, por lo que es fácil consultarlos. Debe configurar un conector en Amazon EMR para que las tablas del metaalmacén de Hive estén disponibles en el clúster. En el siguiente procedimiento se muestra cómo hacerlo:

1. Elija AWS Glue en la consola y cree una tabla basada en sus datos de origen en Amazon S3. Una tabla en el Catálogo de datos de AWS Glue es la definición de metadatos de los datos. En este contexto, tiene sentido crear la tabla manualmente, creando las columnas que desee a partir de los datos de origen. Para obtener más información sobre la creación de tablas en AWS Glue a partir de datos semiestructurados en Amazon S3, consulte Creación de tablas con la consola en la Guía del usuario de AWS Glue.
2. Configure su configuración como parte de la creación de clústeres. Seleccione la pestaña Configuraciones. Las configuraciones son especificaciones opcionales para su clúster. Cuando introduzcas una configuración, añade JSON como en el siguiente ejemplo, en el que se indica a Trino que utilice el catálogo de datos de AWS Glue como su metabastore externo de Hive para los metadatos de las tablas:

   ```
   {
       "classification": "trino-connector-hive",
       "properties": {
           "hive.metastore": "glue"
       }
   }
   ```

   Como alternativa, puedes aplicar las configuraciones en la sección de configuración del software al crear un clúster.

   Además, puede configurar otros tipos de conectores, por ejemplo, para conectarse con Apache Iceberg. Para obtener más información, consulte Uso de un clúster Iceberg con Trino en la Guía de versiones de Amazon EMR. La configuración de ajustes adicionales es opcional.

Para continuar con los pasos de introducción, consulte. Conéctese al nodo principal del clúster de Amazon EMR y ejecute consultas

Crear un clúster con Trino

A continuación, se describen las opciones de configuración correctas al crear un clúster que desee usar con Trino.

> **Important**
>
> Antes de crear el clúster, complete la configuración del catálogo de datos de AWS Glue como su metaalmacén de Hive, que le recomendamos para empezar. Para obtener más información, consulte Uso de un conector Hive para que los datos estén disponibles para su consulta.

1. En la AWS consola, seleccione Amazon EMR de los servicios. Cuando elige Amazon EMR, si tiene clústeres existentes, se muestra su EMR en EC2 los clústeres.
2. Elija Create cluster. Desde aquí, comienza el proceso de creación de un clúster.
3. Asigne un nombre a su clúster y elija una versión de Amazon EMR. Puede elegir la versión más reciente para el tutorial.
4. Elija el paquete Trino, que tiene la aplicación Trino preseleccionada. Los paquetes se configuran para mayor comodidad cuando se conoce con antelación el propósito del clúster. De lo contrario, simplemente seleccione la casilla de verificación de Trino.
5. Para la configuración del clúster, selecciona Grupos de instancias uniformes. Continúe y elimine los grupos de instancias adicionales.
6. Elige un tipo de instancia. Por lo general, le recomendamos que elija un tipo de instancia con al menos 16 GiB de memoria. Además, para el escalado y el aprovisionamiento de clústeres, elige Establecer el tamaño del clúster manualmente.
7. En este punto, establece la configuración de tu metatienda de Hive para que apunte a Glue AWS . Esto se detalla en la sección. Uso de un conector Hive para que los datos estén disponibles para su consulta Complete esto antes de crear el clúster.
8. Elija Create cluster. Puede tardar varios minutos en finalizar.

   Los pasos que aparecen aquí no cubren todos los pasos de configuración en detalle. Encontrará más información sobre la configuración de un clúster en Planificar, configurar y lanzar clústeres de Amazon EMR.

> **Note**
>
> No seleccione Presto y Trino para usarlos en el mismo clúster. No se admite ejecutarlos juntos. También se recomienda que si ejecutas Trino, no ejecutes ninguna otra aplicación en el clúster, como Spark.

Conéctese al nodo principal del clúster de Amazon EMR y ejecute consultas

Aprovisione datos de prueba y configure permisos

Puede probar Amazon EMR con Trino mediante AWS Glue Data Catalog y su metatienda Hive. Los siguientes pasos previos describen cómo configurar los datos de la prueba, si aún no lo ha hecho:

1. Si aún no lo ha hecho, cree una clave SSH para cifrar las comunicaciones.
2. Puede elegir entre varios sistemas de archivos para almacenar datos y archivos de registro. Para empezar, cree un bucket de Amazon S3. Asigne un nombre único al bucket. Al crearlo, especifique la clave de cifrado que creó.

   > **Note**
   >
   > Elija la misma región para crear el depósito de almacenamiento y el clúster de Amazon EMR.

3. Elija el depósito que ha creado. Selecciona Crear carpeta y dale a la carpeta un nombre fácil de recordar. Al crear la carpeta, elija una configuración de seguridad. Puede elegir la configuración de seguridad para la principal o hacer que la configuración de seguridad sea más especializada.
4. Añada los datos de prueba a la carpeta. Para los fines de este tutorial, el uso de un archivo.csv de registros separados por comas funciona bien para completar este caso de uso.
5. Tras añadir datos a un bucket de Amazon S3, configura una tabla en AWS Glue para proporcionar una capa de abstracción para consultar los datos.

Conectar y ejecutar consultas

A continuación, se describe cómo conectarse a un clúster que ejecuta Trino y cómo ejecutarlas en él. Antes de hacerlo, asegúrese de configurar el conector del metaalmacén de Hive, que se describe en el procedimiento anterior, de modo que las tablas del metaalmacén estén visibles.

1. Recomendamos usar EC2 Instance Connect para conectarse al clúster, ya que proporciona una conexión segura. Elija Conectarse al nodo principal mediante SSH en el resumen del clúster. La conexión requiere que el grupo de seguridad tenga una regla de entrada que permita las conexiones a través del puerto 22 a los clientes de la subred. También debe usar el usuario hadoop al conectarse.
2. Inicie la CLI de Trino ejecutando`trino-cli`. Esto le permite ejecutar comandos y consultar datos con Trino.
3. Ejecute `show catalogs;`. Compruebe que el catálogo de colmenas esté en la lista. Esto proporciona una lista de los catálogos disponibles, que contienen almacenes de datos o configuraciones del sistema.
4. Para ver los esquemas disponibles, ejecute. `show schemas in hive;` Desde aquí, puede ejecutar `use` *`schema-name`*`;` e incluir el nombre de su esquema. A continuación, puede correr `show tables;` para enumerar las tablas.
5. Consulta una tabla ejecutando un comando como`SELECT * FROM` *`table-name`*, por ejemplo, usando el nombre de una tabla de tu esquema. Si ya ejecutó la `USE` instrucción para conectarse a un esquema específico, no tiene que usar una notación de dos partes, como*`schema`*. *`table`*.

Configuración de Trino en Amazon EMR

Temas

- Configuración de conectores para Trino
- Monitorización

Configuración de conectores para Trino

Conéctate a AWS Glue como tu metatienda de Hive

Es importante y útil entender que puedes configurar AWS Glue Data Catalog como tu metabastore de Hive cuando ejecutas consultas con Trino. Para obtener información adicional, incluidos los pasos

para configurar un clúster con un metaalmacén de Hive, consulte Uso del catálogo de datos de AWS Glue como metaalmacén de Hive.

Para obtener información sobre la integración de EMR en EKS con AWS Glue, consulte las siguientes prácticas recomendadas, integración de contenedores EMR con Glue. AWS

Conexión a tablas Iceberg al usar Trino con Amazon EMR

Iceberg es un formato de tabla abierto para tablas analíticas. Se creó para que motores como Spark y Trino consultaran macrodatos de las mismas tablas mediante consultas SQL. Incluye funciones como aislar las lecturas y escrituras de datos, para que el lector pueda evitar consultar datos que estén parcialmente actualizados, por ejemplo. También es compatible con funciones de estado, como las instantáneas. Proporciona una capa de abstracción mediante el uso de metadatos y archivos de manifiesto. Describen el esquema de la tabla y facilitan la consulta de datos sin tener que conocer muchos detalles sobre su formato u organización. Cuando está conectado, puede leer los datos de las tablas, actualizar los datos o escribir nuevos datos en los archivos subyacentes.

Hay un taller disponible en el que se muestra cómo configurar tablas Iceberg con Amazon EMR y AWS Glue. Para obtener más información, consulte el taller de análisis: configuración y uso de tablas Iceberg de Apache en su lago de datos.

Conexión con los clientes

Puede conectarse con Trino mediante un controlador JDBC disponible. Para obtener más información, consulte el controlador JDBC en la documentación de Trino.

Monitorización

Puede supervisar los clústeres de Amazon EMR a través del. AWS Management Console Para obtener más información, consulte Visualización y monitorización de un clúster de Amazon EMR a medida que realiza su trabajo. Amazon EMR también envía sus métricas de supervisión a. Amazon CloudWatch Para obtener más información sobre la monitorización de un clúster de Amazon EMR, consulte Amazon CloudWatch Eventos y métricas de Amazon EMR.

Prácticas recomendadas para Trino en Amazon EMR

La arquitectura de Trino está diseñada para realizar consultas SQL rápidas y distribuidas en grandes conjuntos de datos de varias fuentes de datos, siguiendo un modelo de coordinador-trabajador, en el que cada componente desempeña una función especializada en la ejecución de las consultas. Hay

algunas áreas o categorías en las que puede centrarse para configurar su clúster de Amazon EMR que ejecuta Trino para obtener el mejor rendimiento. Estos incluyen los siguientes:

- Ajustar los ajustes de configuración del clúster para optimizar la memoria.
- Optimización de los ajustes para el particionamiento y la distribución de datos.
- Uso del filtrado dinámico para reducir el recuento de resultados de consultas.

Algunos de estos ajustes se ajustan automáticamente cuando utiliza Trino con Amazon EMR. Otros pueden configurarse manualmente a través de la consola o mediante comandos de la CLI. Los temas de esta sección le ayudan a configurar los datos y su clúster de forma óptima.

Temas

- Áreas clave de enfoque para la mejora del rendimiento
- Recopile y utilice estadísticas de tablas
- Desafíos habituales a la hora de escalar las cargas de trabajo de Trino

Áreas clave de enfoque para la mejora del rendimiento

Trino maximiza el paralelismo de las consultas y la optimización de la memoria. Esta arquitectura proporciona flexibilidad al permitirle consultar múltiples y variadas fuentes de datos y, al mismo tiempo, escalar de manera eficiente. Las áreas clave de la mejora del rendimiento en Trino incluyen las que se enumeran a continuación.

Optimizada de la memoria

La administración de la memoria en Trino es fundamental para lograr un alto rendimiento y estabilidad, especialmente cuando se ejecutan consultas grandes y complejas. Trino utiliza un modelo de memoria distribuida. En este modelo, la memoria se asigna entre los nodos de trabajo para procesar tareas, agregaciones, uniones y otras operaciones. La siguiente lista presenta un conjunto de estos ajustes:

- query.max-memory: establece la memoria máxima disponible para una sola consulta en todo el clúster. Este es un límite estricto; si una consulta supera esta memoria, fallará.
- consulta. max-memory-per-node — Define la memoria máxima que puede consumir una consulta en cada nodo de trabajo. Si se establece esto, se garantiza que ninguna consulta monopolice los recursos de ningún trabajador.

- Tamaño del montón de JVM: configurado a nivel de JVM, establece el tamaño máximo del montón para el proceso del servidor Trino en cada nodo. Por lo general, este valor debe ser mayor que las configuraciones relacionadas con la memoria (es la suma de las consultas). max-memory-per-nodey memoria. heap-headroom-per-node) en Trino para evitar que el sistema se quede sin memoria a nivel de la JVM.
- memoria. heap-headroom-per-node — Especifica la cantidad de memoria del búfer que se debe dejar fuera del tamaño del montón de la JVM para operaciones que no sean de consulta. Esto es crucial para garantizar una sobrecarga suficiente para las operaciones internas y la recolección de basura.

Filtrado dinámico

El filtrado dinámico de Trino es una técnica de optimización que mejora el rendimiento de las consultas al reducir la cantidad de datos procesados, especialmente durante las uniones. Aplica condiciones de filtrado de forma dinámica para limitar los datos escaneados por un lado de una combinación, en función de los datos que se ven en el otro lado, lo que resulta especialmente útil en consultas en las que un lado de la combinación es muy selectivo (lo que significa que contiene un pequeño subconjunto de datos). Está habilitada de forma predeterminada en Amazon EMR. A continuación se muestra una consulta de ejemplo:

```
SELECT orders.order_id, orders.total_amount
FROM orders
JOIN customers ON orders.customer_id = customers.customer_id
WHERE customers.country = 'France';
```

Sin un filtrado dinámico, Trino escanea toda la tabla de pedidos de una sola vez, aunque solo sea relevante un pequeño subconjunto de clientes (los de Francia). Este enfoque lee todas las filas de la tabla de pedidos, lo que se traduce en altos costes de procesamiento y E/S. Con el filtrado dinámico, Trino escanea inicialmente la tabla de clientes más pequeña, recupera los valores customs_id solo de los clientes de Francia y, a continuación, aplica este subconjunto como filtro a los pedidos. Esto significa que solo se escanean las filas relevantes de los pedidos (aquellas con un custome_id que coincide con el subconjunto filtrado), lo que reduce considerablemente los registros procesados.

Derramar al disco

En Trino, el almacenamiento de discos permite descargar al disco los resultados de las consultas intermedias, lo que permite completar las consultas que consumen mucha

memoria, incluso si superan los límites de memoria establecidos por o. `query_max_memory` `query_max_memory_per_node` De forma predeterminada, Trino impone estos límites para garantizar una asignación de memoria equitativa y evitar que el clúster se bloquee. Sin embargo, cuando una consulta grande supera estos límites, corre el riesgo de ser finalizada. La dispersión de discos soluciona este problema mediante el uso`revocable memory`, lo que permite que una consulta tome prestada memoria adicional que se puede revocar si se necesitan recursos en otros lugares. Cuando se anula la memoria, los datos intermedios se acumulan en el disco, lo que permite que las consultas continúen procesándose sin sobrepasar los límites de memoria. Tenga en cuenta que una consulta que se ve obligada a extenderse al disco puede tener un tiempo de ejecución más prolongado, por lo que está deshabilitada de forma predeterminada. Para habilitar el derrame en Amazon EMR, utilice la siguiente configuración:

- `spill-enabled=true`— Permite que el disco se derrame cuando el uso de memoria supera los umbrales disponibles.
- `spill-paths`— Define los directorios donde se almacenan los datos derramados,. `spill-paths=/mnt/spill`

Recopile y utilice estadísticas de tablas

La recopilación de estadísticas de tablas permite al optimizador basado en costes de Trino tomar decisiones informadas sobre las órdenes de unión, el uso de filtros y la reducción de particiones, lo que se traduce en un mejor rendimiento.

Puede utilizar el `ANALYZE` comando para recopilar estadísticas para las tablas Hive o Iceberg:

```
ANALYZE sales;
```

Recopilar estadísticas en tablas anchas puede resultar agotador para los recursos. Se recomienda especificar un subconjunto de columnas que se utilicen en las uniones, los filtros o las operaciones de agrupación.

Este es otro comando útil. Muestra las estadísticas actuales de una tabla para comprobar si las estadísticas están actualizadas.

```
show stats for table_name;
```

Desafíos habituales a la hora de escalar las cargas de trabajo de Trino

Los principales beneficios de usar Amazon S3 con Trino son la capacidad de S3 para escalar grandes volúmenes de datos y la rentabilidad de S3. Sin embargo, cuando consulta grandes volúmenes de datos, en ocasiones puede producirse una serie de problemas de rendimiento relacionados con el rendimiento. Estos pueden deberse a la forma en que se almacenan los datos, a los ajustes de configuración que restringen el buen rendimiento o a otros motivos. Cuando se producen estos problemas, hay medidas eficaces que puede tomar para evitarlos o mitigarlos.

Esta sección comienza con una lista de optimizaciones generales que puede implementar para aumentar el rendimiento de las consultas en grandes volúmenes de datos. A continuación, se detallan los problemas comunes y se proporcionan las mitigaciones para cada uno de ellos.

Este tema proviene de la siguiente presentación de la conferencia: Acelere el rendimiento a escala: prácticas recomendadas para Trino con Amazon S3.

Optimizada del diseño de datos para grandes conjuntos de datos

Los cuellos de botella en el rendimiento no son infrecuentes cuando se consultan conjuntos de datos de gran tamaño. Sin embargo, existen prácticas recomendadas que puede implementar para tener una ventaja inicial a la hora de utilizar Trino para consultar datos en Amazon S3. Estos incluyen los siguientes:

- Particionamiento: particionar significa organizar los datos en una jerarquía y almacenar los datos relacionados juntos, en función de los atributos relacionados. La partición permite que las consultas no tengan que escanear tantos datos irrelevantes y mejora el rendimiento de las consultas. Puede usar varias estrategias de partición, como organizar los datos de origen con prefijos, específicamente por rangos de fechas, regiones u otros atributos. Para obtener información más detallada sobre cómo particionar los datos en Amazon S3 para aumentar el rendimiento, consulte la entrada del blog Comience a administrar particiones para las tablas de Amazon S3 respaldadas por el catálogo de datos de AWS Glue o la publicación Los 10 mejores consejos de ajuste del rendimiento para Amazon Athena.
- Agrupación: agrupar datos relacionados en archivos comunes. Por ejemplo, si consultas datos según una región geográfica, como un estado, puedes mejorar el rendimiento de las consultas agrupando todos los datos de un estado concreto en el mismo archivo o grupo de archivos. Para que esto funcione mejor, basa la agrupación en un atributo de datos con una cardinalidad alta, como un estado o una provincia, por ejemplo. Además, puedes tener en cuenta tus patrones de

consulta. Un ejemplo de esto podría consistir en agrupar los datos de California y Oregón, si las consultas suelen leer datos de esos estados juntos.

- Administración de prefijos S3: puede utilizar los prefijos de Amazon S3 para implementar una estrategia de particionamiento. Si utilizas un único prefijo para un bucket de Amazon S3, como una fecha concreta, por ejemplo, esto puede provocar un número elevado de solicitudes y provocar un error HTTP 503. Se recomienda utilizar prefijos para añadir condiciones adicionales y organizar los datos de origen de forma más eficaz. Para obtener más información, consulte Organizar objetos con prefijos en la documentación de Amazon S3. El siguiente breve ejemplo muestra un prefijo que mejora el rendimiento de las solicitudes:. `s3://bucket/country=US/dt=2024-06-13` En este ejemplo, tanto el país como la fecha están incluidos en el prefijo, lo que resulta en menos lecturas que en un caso en el que el prefijo incluye solo la fecha.

  La mitigación de los errores 503 del HTTP se analiza con más detalle en la sección de ralentización del HTTP que aparece a continuación en este tema.

- Optimización del tamaño de los datos: puede ejecutar el comando OPTIMIZE para establecer una configuración que permita que las consultas tengan un mejor rendimiento. Para ejecutarlo en tablas externas de Hive, sigue estos pasos:

  - `OPTIMIZE`Utilícelo con el siguiente parámetro:`hive.non-managed-table-writes-enabled=true`. Para obtener más información sobre esta propiedad, consulte las propiedades de configuración general de Hive.
  - Defina el siguiente parámetro de sesión: `SET SESSION` *`catalog`*`.non_transactional_optimize_enabled=true`
  - Ejecute el `OPTIMIZE` comando:`ALTER TABLE` *`catalog`*`.`*`schema`*`.`*`table`* `EXECUTE optimize(file_size_threshold => '128MB')`. En este caso, `file_size_threshold` es de 100 MB por defecto. Si se aumenta este umbral, como se muestra en el ejemplo, se fusionarán los archivos de menos de 128 MB.

- Configurar los reintentos: puede aumentar el límite de reintentos, lo que puede mitigar la posibilidad de que se produzcan errores en el HTTP 503, configurando lo siguiente:. `s3.max-error-retries` Esto se aplica cuando utiliza la TrinoFileSystem API y la versión 449 de Trino o posterior. Por otro lado, en el caso de que utilice Amazon EMR con Trino, utilizará EMRFS para acceder a Amazon S3. Con EMRFS, puede aumentar el número de retiros cambiando el parámetro. `fs.s3.maxRetries`

- Elija una clase de almacenamiento de Amazon S3: elegir la clase de almacenamiento adecuada para los datos en diferentes momentos de su ciclo de vida puede mejorar tanto el rendimiento como el costo, en función de sus requisitos para recopilaciones de datos específicas. Para obtener

más información, consulte Entender y administrar clases de almacenamiento de Amazon S3 en la documentación de Amazon S3.

- Migrar a Iceberg: otra solución para mitigar los problemas de rendimiento, específicamente relacionados con la ejecución de consultas en archivos pequeños, consiste en migrar a las tablas de Iceberg. Iceberg tiene funciones que permiten gestionar bien los archivos pequeños.
- Utilice la compactación automática de datos: si utiliza tablas Iceberg, la compactación automática de datos con el catálogo de datos de AWS Glue puede optimizar el tamaño de los datos y mejorar el rendimiento de las consultas.

Desafíos habituales cuando se consultan conjuntos de datos de gran tamaño

En esta sección se incluye una colección de problemas comunes que pueden producirse al acumular un conjunto de datos de gran tamaño en Amazon S3 y consultarlo con Trino. En cada sección se muestran formas de resolver el problema o reducir su impacto en las consultas. Cada uno de los problemas descritos en las siguientes secciones se ha reproducido y probado con un conector Hive.

Análisis de datos de gran tamaño

Cuando la consulta tiene que escanear conjuntos de datos de gran tamaño, se pueden producir problemas como un rendimiento lento de las consultas y un mayor coste de almacenamiento. Los grandes volúmenes de datos pueden deberse a un rápido crecimiento de los datos o a una planificación que no implica el traslado de los datos heredados en un plazo adecuado. Esto puede provocar consultas más lentas.

Para mitigar los impactos en el rendimiento derivados del escaneo de grandes conjuntos de datos, le recomendamos que utilice particiones y agrupaciones:

- La partición agrupa los datos relacionados en función de sus atributos. El uso eficaz de las particiones puede mejorar considerablemente el rendimiento de las consultas.
- El agrupamiento se refiere a agrupar los datos en archivos o depósitos según columnas de datos específicas y relacionadas. Por lo general, agrupar en grupos significa mantener juntos físicamente los archivos de datos de origen relacionados.

Para ilustrar cómo puede funcionar la mitigación en escaneos de datos de gran tamaño, supongamos que almacena y consulta datos que tienen registros con un atributo de estado, que puede asignarse a California o Alaska, y que este atributo de estado es una de las condiciones de consulta. Puede mejorar el rendimiento de las consultas almacenando los datos de cada estado

en un depósito de S3 independiente o dividiendo los datos en función del estado con un prefijo S3. Esta división y agrupamiento también pueden mejorar el rendimiento si se basan en una columna adicional, como un atributo de fecha, por ejemplo.

> **Note**
>
> Si una columna tiene una cardinalidad alta y desea utilizarla para agrupar datos, le recomendamos que utilice la agrupación en grupos en este caso. Por otro lado, por lo general, las claves de partición deberían tener una cardinalidad más baja.

Uso de varios tipos de almacenamiento S3

Por lo general, usted elige los tipos de almacenamiento en función del rendimiento, el acceso a los datos, la resiliencia y los requisitos de costo de sus cargas de trabajo. Puede haber compensaciones entre el costo y el rendimiento. Es importante elegir la clase de almacenamiento de Amazon S3 adecuada que se adapte a sus patrones de acceso a los datos. Hay dos patrones de acceso principales:

- Datos a los que se accede de forma conocida o predecible. Por lo general, si tiene datos a los que se accede con poca frecuencia, S3 Standard IA puede ser una buena opción, ya que ayuda a reducir los costos. Si ha accedido a los datos con frecuencia, S3 Standard es la mejor opción para acceder con Amazon EMR y Trino.
- Datos a los que se accede de forma desconocida o impredecible. Esto puede requerir el uso de otras clases de almacenamiento de Amazon S3. Existen compensaciones entre las clases de almacenamiento de S3. Estos incluyen la latencia, el costo de almacenamiento y la disponibilidad. Puede elegir un tipo de almacenamiento de S3 adecuado, en función de sus cargas de trabajo y patrones de acceso. Para obtener descripciones de las ventajas de cada clase, consulte Clases de almacenamiento de Amazon S3.

Uso de la compactación

También puede utilizar la compactación automática de Iceberg, si utiliza tablas de Iceberg, lo que da como resultado tamaños de archivo más óptimos para aumentar la eficacia de las consultas. Para obtener más información, consulte El catálogo de datos de AWS Glue ahora admite la compactación automática de tablas de Apache Iceberg.

Errores de ralentización de HTTP

Esto ocurre cuando la tasa de solicitudes supera un umbral preconfigurado en un prefijo de Amazon S3. El error HTTP que se produce con más frecuencia cuando se alcanza este estado es el siguiente: Error 503: reduzca la tasa de solicitudes. El origen de este problema puede deberse a la presencia de una gran cantidad de archivos pequeños, debido a la cantidad de divisiones que se deben crear para poder leer los datos. Hay un par de maneras de mitigar el problema:

- Aumente el límite de reintentos para las solicitudes de Amazon S3 en Trino. Esto está configurado para los EMRFS que se utilizan `fs.s3.maxretries` en Trino 449.
- Optimice los tamaños de los archivos, lo que también puede reducir la tasa de solicitudes.

Para obtener más información sobre cómo Trino determina el número de divisiones en un conjunto de datos que se van a consultar, consulte las propiedades de configuración del ajuste del rendimiento en la documentación del conector de Hive.

Dificultad para consultar archivos pequeños

La consulta de muchos archivos pequeños puede provocar una sobrecarga de E/S excesiva, debido a la gran cantidad de solicitudes GET y LIST, y, posteriormente, afectar negativamente al rendimiento de las consultas. La optimización del tamaño de los archivos puede mejorar el rendimiento de las consultas. Existen varias formas de realizar esta operación:

- Consolide los datos en menos archivos de gran tamaño. (Por lo general, se recomienda mantener el tamaño de los archivos en torno a los 128 MB). Puede hacerlo con herramientas al ingerir datos, como en una canalización de ETL, o puede consolidar los datos manualmente. Si estas soluciones no están disponibles, las opciones restantes podrían ser más adecuadas para usted.
- Ejecute el comando `OPTIMIZE`.
- Establezca el parámetro `SESSION`.

Tenga en cuenta que Iceberg tiene una función disponible para combinar archivos pequeños en archivos más grandes, que es la compactación automática. Funciona con archivos gestionados con el catálogo de datos de AWS Glue. Para obtener más información, consulte El catálogo de datos de AWS Glue ahora admite la compactación automática de tablas de Apache Iceberg.

Consultas que incluyen datos que no son necesarios

Es habitual que los datos crezcan, por lo que es imprescindible realizar un seguimiento de los patrones de acceso a los datos y moverlos de forma adecuada a medida que envejecen o se vuelven irrelevantes. Esto se debe a que, a medida que aumentan los datos, el rendimiento de las consultas puede degradarse con el tiempo, principalmente debido al enorme volumen de datos que hay que analizar cuando se ejecuta una consulta. Amazon S3 y otros servicios ofrecen orientación para la migración del ciclo de vida de los datos, en la que se muestran las estrategias para mover los datos a diferentes ubicaciones de almacenamiento a medida que se enfrían. Hacer esto también supone una ventaja en cuanto a costes de almacenamiento.

Además de la migración de datos, puede utilizar otras estrategias, como eliminar los datos de origen que no sean relevantes para las consultas que está ejecutando. Esto puede llevar algo de trabajo, ya que podría implicar cambiar el esquema de los datos de origen. Sin embargo, su resultado positivo es reducir el volumen de datos y agilizar las consultas. Para obtener más información, consulte Gestión del ciclo de vida de objetos.

Consideraciones sobre Trino

Tenga en cuenta lo siguiente cuando ejecute Trino en Amazon EMR.

Propiedades de la implementación de Trino no configurables

En la siguiente tabla, se muestran las distintas opciones de configuración de los `properties` archivos Trino.

Archivos	Configurable
`log.properties`	Trino: versiones 6.1.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `prestosql-log` o `trino-log` .
`config.properties`	Trino: versiones 6.1.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `prestosql-config` o `trino-config` .

Archivos	Configurable
`hive.properties`	Trino: versiones 6.1.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `prestosql-connector-hive` o `trino-connector-hive` .
`node.properties`	Trino: versiones 6.1.0 y posteriores configurables de Amazon EMR. Utilice la clasificación de configuración `prestosql-node` o `trino-node` .
`jvm.config`	No se puede configurar.

Consideraciones adicionales

- Para la versión 6.1.0 y posteriores de Trino, Amazon EMR configura automáticamente una clave secreta compartida para una comunicación interna segura entre los nodos del clúster. No necesita llevar a cabo ninguna configuración adicional para habilitar esta característica de seguridad y puede anular la configuración con su propia clave secreta. Para obtener información sobre la autenticación interna de Trino, consulte la documentación de Trino 353: Comunicación interna segura.

Historial de lanzamientos de Trino

Las secciones de notas de la versión detallan los cambios y actualizaciones de una versión específica de Trino en Amazon EMR.

Notas de la versión de Trino por versión

- Amazon EMR 7.6.0: notas de la versión de Trino
- Amazon EMR 7.3.0: notas de la versión de Trino
- Amazon EMR 6.9.0: notas de la versión de Trino

Amazon EMR 7.6.0: notas de la versión de Trino

Amazon EMR 7.6.0: nuevas características de Trino

- Para admitir consultas de ejecución prolongada, Trino ahora incluye un mecanismo de ejecución tolerante a errores. La ejecución tolerante a errores mitiga los errores de las consultas al volver a intentar las consultas con errores o las tareas que las componen.

Amazon EMR 7.6.0: cambios en Trino

Amazon EMR 7.6.0: cambios en Trino

Tipo	Descripción
Upgrade	Actualización de Trino a 457

Amazon EMR 7.3.0: notas de la versión de Trino

Amazon EMR 7.3.0: cambios en Trino

- Esta versión actualiza Trino de la versión 436 a la 442.
- Esta versión redirige las consultas de Hudi al nuevo corrector de Hudi. El antiguo conector Hive ya no puede leer las tablas Hudi. Nota
- Esta versión elimina el módulo Rubix de Amazon EMR porque ahora está en desuso desde el código abierto.
- Esta versión elimina el modo heredado de la propiedad `hive.security`. El valor predeterminado es `allow-all`.

Amazon EMR 6.9.0: notas de la versión de Trino

Amazon EMR 6.9.0: nuevas características de Trino

- Para admitir consultas de ejecución prolongada, Trino ahora incluye un mecanismo de ejecución tolerante a errores. La ejecución tolerante a errores mitiga los errores de las consultas al volver a intentar las consultas con errores o las tareas que las componen.

Amazon EMR 6.9.0: cambios en Trino

Amazon EMR 6.9.0: cambios en Trino

Tipo	Descripción
Upgrade	Actualización de Trino a 398
Upgrade	Compatibilidad con Hadoop 3.3.3
Característica	Compatibilidad con Tardigrade: se agrega compatibilidad con el intercambio en cadena en HDFS y Amazon S3.
Corrección de errores	Cuando se usa Trino Iceberg y el catálogo de Glue esté activado, se evita agregar la URL de metaalmacén en `iceberg.properties`.

Amazon EMR 6.9.0: problemas conocidos de Trino

- En el caso de la versión 6.9.0 de Amazon EMR, Trino no funciona en clústeres habilitados para Apache Ranger. Si tiene que utilizar Trino con Ranger, contacte con Soporte.

Apache Zeppelin

Usa Apache Zeppelin como cuaderno para la exploración interactiva de datos. Para obtener más información sobre Zeppelin, consulte https://zeppelin.apache.org/. Zeppelin se incluye en la versión 5.0.0 de Amazon EMR y posteriores. Las versiones anteriores incluyen Zeppelin como una aplicación de entorno aislado. Para obtener más información, consulte Versiones de lanzamiento de Amazon EMR 4.x.

Para acceder a la interfaz web de Zeppelin, configure un túnel de SSH en el nodo principal y una conexión de proxy. Para obtener más información, consulte Ver interfaces web alojadas en clústeres de EMR.

En la siguiente tabla, se muestra la versión de Zeppelin incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes que Amazon EMR instala con Zeppelin.

Para ver la versión de los componentes instalados con Zeppelin en esta versión, consulte Versiones de los componentes de la versión 7.9.0.

Información de la versión de Zeppelin para emr-7.9.0

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-7.9.0	Zeppelin 0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

En la tabla siguiente, se muestra la versión de Zeppelin que se incluye en la última versión de la serie Amazon EMR 6.x, junto con los componentes que Amazon EMR instala con Zeppelin.

Para ver la versión de los componentes instalados con Zeppelin en esta versión, consulte Versiones de los componentes de la versión 6.15.0.

Información de la versión de Zeppelin para emr-6.15.0

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.15.0	Zeppelin 0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

En la tabla siguiente, se muestra la versión de Zeppelin que se incluye en la última versión de la serie Amazon EMR 5.x, junto con los componentes que Amazon EMR instala con Zeppelin.

Para ver la versión de los componentes instalados con Zeppelin en esta versión, consulte Versiones del componente de la versión 5.36.2.

Información de la versión de Zeppelin para emr-5.36.2

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.36.2	Zeppelin 0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
		client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Temas

- Consideraciones sobre el uso de Zeppelin en Amazon EMR
- Historial de versiones de Zeppelin

Consideraciones sobre el uso de Zeppelin en Amazon EMR

- Conecte a Zeppelin utilizando el mismo método de configuración de un túnel de SSH que utiliza para conectar a otros servidores web en el nodo principal. El servidor de Zeppelin se encuentra en el puerto 8890.
- Zeppelin en la versión de lanzamiento 5.0.0 de Amazon EMR y posteriores admite la autenticación de Shiro.
- Zeppelin en la versión 5.8.0 y posteriores de Amazon EMR admite el uso del Catálogo de datos de AWS Glue como metaalmacén de Spark SQL. Para más información, consulte Uso del Catálogo de datos de AWS Glue como metaalmacén de Spark SQL.
- Zeppelin no utiliza algunos de los ajustes definidos en el archivo de configuración `spark-defaults.conf` de su clúster, aunque envía instrucciones a YARN para asignar ejecutores dinámicamente si ha establecido `spark.dynamicAllocation.enabled` como `true`. Debe definir los ajustes del ejecutor (por ejemplo, memoria y núcleos) en la pestaña Interpreter (Intérprete) de Zeppelin y, a continuación, reiniciar el intérprete para utilizarlos.

- Las versiones 6.10.0 y posteriores de Amazon EMR admiten la integración de Apache Zeppelin con Apache Flink. Para obtener más información, consulte Trabajar con trabajos de Flink de Zeppelin en Amazon EMR.
- Zeppelin en Amazon EMR no admite el intérprete de SparkR.

Historial de versiones de Zeppelin

En la siguiente tabla, se muestra la versión de Zeppelin que se incluye en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

Información de la versión de Zeppelin

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-7.9.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.8.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.7.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-7.6.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.5.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-7.4.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.3.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-7.2.0	0.10.1-amz	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.2	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-7.1.0	0.10.1-amz	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.0.0	0.10.1-amz	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.15.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.14.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.13.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.12.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.11.1	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.11.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.10.1	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.10.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.9.1	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.9.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.8.1	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.8.0	0.10.1-amz	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.7.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.1	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.36.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.6.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.35.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.5.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.4.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.3.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.3.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.2.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.2.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.1.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.1.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.0.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-6.0.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.34.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.33.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.33.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.32.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.32.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.31.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.31.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.30.2	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.30.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.30.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.29.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.28.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.28.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.27.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.27.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.26.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.25.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.24.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.24.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.23.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.23.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.22.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.2	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.21.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.20.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.20.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.19.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.19.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.18.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.18.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.17.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.17.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.17.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.16.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.16.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.15.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.15.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.14.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.13.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.13.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.12.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.12.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.11.4	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.11.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.11.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.10.1	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.10.0	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.9.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.9.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.3	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.8.2	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.8.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.7.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.7.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.6.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.6.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.4	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.5.3	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.2	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.5.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.4.1	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.4.0	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.3.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.3.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.3.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.3	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.2.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.2.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.1.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.1.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.0.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Etiqueta de versión de Amazon EMR	Versión de Zeppelin	Componentes instalados con Zeppelin
emr-5.0.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Apache ZooKeeper

Apache ZooKeeper es un servicio centralizado para mantener información de configuración, nomenclatura, proporcionar sincronización distribuida y proporcionar servicios de grupo. Para obtener más información al respecto ZooKeeper, consulte http://zookeeper.apache.org/.

En la tabla siguiente, se muestra la versión ZooKeeper incluida en la última versión de la serie 7.x de Amazon EMR, junto con los componentes con los que Amazon EMR instala. ZooKeeper

Para ver la versión de los componentes instalados con ZooKeeper esta versión, consulte Versiones de los componentes de la versión 7.9.0.

ZooKeeper Información de la versión de emr-7.9.0

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr7.9.0	ZooKeeper 3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

En la tabla siguiente, se muestra la versión ZooKeeper incluida en la última versión de la serie Amazon EMR 6.x, junto con los componentes con los que se instala Amazon EMR, junto con los componentes con los que se instala Amazon EMR. ZooKeeper

Para ver la versión de los componentes instalados con ZooKeeper esta versión, consulte Versiones de los componentes de la versión 6.15.0.

ZooKeeper Información de la versión de emr-6.15.0

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.15.0	ZooKeeper 3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

En la tabla siguiente, se muestra la versión ZooKeeper incluida en la última versión de la serie Amazon EMR 5.x, junto con los componentes con los que Amazon EMR instala. ZooKeeper

Para ver la versión de los componentes instalados con ZooKeeper esta versión, consulte Versiones de los componentes de la versión 5.36.2.

ZooKeeper Información de la versión de emr-5.36.2

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.36.2	ZooKeeper 3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Temas

- ZooKeeper historial de versiones
- ZooKeeper notas de la versión

ZooKeeper historial de versiones

En la tabla siguiente, se muestra la versión ZooKeeper incluida en cada versión de lanzamiento de Amazon EMR, junto con los componentes instalados con la aplicación. Para ver las versiones de los componentes de cada versión, consulte la sección Versiones de los componentes en la versión correspondiente de Versiones de Amazon EMR 7.x, Versiones de lanzamiento de Amazon EMR 6.x o Versiones de lanzamiento de Amazon EMR 5.x.

ZooKeeper información sobre la versión

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr7.9.0	3,9,3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr7.8.0	3,9,3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
		timeline-server, zookeeper-client, zookeeper-server
emr-7.7.0	3,9,2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.6.0	3,9,2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-7.5.0	3,9,2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.4.0	3,9,2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.3.0	3,9,1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-7.2.0	3,9,1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.1.0	3,9,1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-7.0.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.15.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.14.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.13.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.12.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.11.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.10.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.10.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.9.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.9.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.8.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.7.0	3,5,7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.36.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.6.0	3,5,7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.35.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.5.0	3,5,7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.4.0	3,5,7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.3.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.2.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.2.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.1.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.1.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-6.0.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.34.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.33.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.32.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.32.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.31.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.31.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.30.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.29.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.28.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.28.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.27.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.26.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.25.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.24.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.24.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.23.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.22.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.2	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.21.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.20.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.19.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.19.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.18.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.18.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.2	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.17.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.16.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.15.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.15.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.14.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.13.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.13.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.12.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.11.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.11.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.10.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.9.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.9.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.8.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.8.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.7.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.7.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.6.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.6.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.5.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.5.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.5.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.4.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.4.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.3.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.2.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.2.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.1.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.1.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.0.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Etiqueta de versión de Amazon EMR	ZooKeeper Versión	Componentes instalados con ZooKeeper
emr-5.0.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

ZooKeeper notas de la versión

Consulte las siguientes secciones para ver las notas completas de la versión.

Amazon EMR 7.8.0: notas de la versión ZooKeeper

Amazon EMR 7.8.0: cambios ZooKeeper

Actualización: ZooKeeper se ha actualizado a la versión 3.9.3.

Amazon EMR 7.5.0: notas de la versión ZooKeeper

Amazon EMR 7.5.0: cambios ZooKeeper

- A partir de EMR-7.5.0, ZooKeeper está configurado en Java 17 de forma predeterminada en tiempo de ejecución. Para usar una versión diferente para el entorno de ejecución de Java, anule la configuración de JVM del `zookeeper-server` archivo y ajústela a la versión deseada. `JAVA_HOME`

Amazon EMR 7.4.0: notas de la versión ZooKeeper

Tipo	Descripción
Upgrade	La versión de ZooKeeper se ha actualizado a la versión 3.9.2.
Mejora	ZOOKEEPER-4778: parchee jetty, netty y logback para eliminar las vulnerabilidades de alta gravedad.
Mejora	ZOOKEEPER-4799: corrige el problema CVE-2024-23944.
Corrección de errores	TLSv1Se ha añadido el soporte para ZOOKEEPER-4415:3.

Amazon EMR 7.4.0: características ZooKeeper

- Los clústeres con una configuración de seguridad de cifrado en tránsito tendrán el TLS activado de forma predeterminada para la comunicación por quórum y el servidor de administración de Zookeeper. Además, el puerto seguro TLS está expuesto en 2281 para la comunicación del cliente con Zookeeper.

Conectores y utilidades

Amazon EMR proporciona varios conectores y utilidades para acceder a otros AWS servicios como fuentes de datos. Normalmente, es posible acceder a los datos de estos servicios dentro de un programa. Por ejemplo, puede especificar una transmisión de Kinesis en una consulta de Hive, un script de Pig o una MapReduce aplicación y, a continuación, operar con esos datos.

Temas

- Exportación, importación, consulta y unión de tablas en DynamoDB mediante Amazon EMR
- Kinesis
- S3 (s3-dist-cp) DistCp
- Limpiando después de DistCp trabajos fallidos en S3

Exportación, importación, consulta y unión de tablas en DynamoDB mediante Amazon EMR

> **Note**
>
> El conector Amazon EMR-DynamoDB es de código abierto en. GitHub Para obtener más información, consulte https://github.com/awslabs/emr-dynamodb-connector.

DynamoDB es un servicio de bases de datos NoSQL totalmente administrado que proporciona un rendimiento rápido y predecible, así como una perfecta escalabilidad. Los desarrolladores pueden crear una tabla de base de datos y ampliar su tráfico de solicitudes o almacenamiento sin límite. DynamoDB distribuye automáticamente los datos y el tráfico para la tabla entre un número de servidores suficiente para administrar la capacidad de solicitudes que el cliente especifique y la cantidad de datos almacenados, manteniendo al mismo tiempo un rendimiento uniforme y rápido. Mediante Amazon EMR y Hive puede procesar de forma rápida y eficaz grandes cantidades de datos, como, por ejemplo, datos almacenados en DynamoDB. Para obtener más información sobre DynamoDB, consulte la Guía para desarrolladores de Amazon DynamoDB.

Apache Hive es una capa de software que puede utilizar para consultar clústeres de MapReduce utilizando un lenguaje de consultas de tipo SQL simplificado denominado HiveQL. Se ejecuta sobre la arquitectura de Hadoop. Para obtener más información acerca de Hive y HiveQL, consulte el

HiveQL Language Manual. Para obtener más información sobre Hive y Amazon EMR, consulte Apache Hive.

Puede utilizar Amazon EMR con una versión personalizada de Hive que incluye conectividad con DynamoDB para realizar operaciones con los datos almacenados en DynamoDB:

- Cargar datos de DynamoDB en Sistema de archivos distribuido de Hadoop (HDFS) y utilizarlos como entrada en un clúster de Amazon EMR.
- Consultar datos de DynamoDB en directo utilizando instrucciones similares a SQL (HiveQL).
- Unir datos almacenados en DynamoDB y exportarlos o consultar respecto a los datos unidos.
- Exportación de datos almacenados en DynamoDB a Amazon S3.
- Importación de datos almacenados en Amazon S3 a DynamoDB.

> **Note**
>
> - El conector Amazon EMR-DynamoDB no es compatible con clústeres configurados para utilizar la autenticación Kerberos.

Para realizar cada una de las siguientes tareas, lanzará un clúster de Amazon EMR, especificará la ubicación de los datos en DynamoDB y emitirá comandos de Hive para manipular los datos en DynamoDB.

Hay varias formas de lanzar un clúster de Amazon EMR: puede utilizar la consola de Amazon EMR, la interfaz de línea de comandos (CLI) o puede programar el clúster mediante un SDK AWS o la API de Amazon EMR. También puede elegir si desea ejecutar un clúster de Hive de forma interactiva o a partir de un script. En esta sección le mostraremos cómo lanzar un clúster de Hive interactivo desde la consola de Amazon EMR y en la CLI.

Utilizar Hive de forma interactiva es una forma fantástica de probar el rendimiento de las consultas y ajustar su aplicación. Después de que haya establecido un conjunto de comandos de Hive que se ejecutarán de forma periódica, considere la posibilidad de crear un script de Hive que Amazon EMR puede ejecutar por usted.

> **Warning**
>
> Las operaciones de lectura o escritura de Amazon EMR en una tabla de DynamoDB se tienen en cuenta para el rendimiento aprovisionado establecido, por lo que es posible

que aumente la frecuencia de las excepciones de rendimiento aprovisionado. En el caso de solicitudes de gran tamaño, Amazon EMR implementa los reintentos con retroceso exponencial para administrar la carga de solicitudes en la tabla de DynamoDB. La ejecución de tareas de Amazon EMR de forma simultánea con otro tráfico podría provocar que supere el nivel de rendimiento aprovisionado asignado. Puedes monitorizarlo comprobando la ThrottleRequestsmétrica en Amazon CloudWatch. Si la carga de solicitudes es demasiado elevada, puede volver a lanzar el clúster y definir Ajuste de porcentaje de lectura o Configuración de porcentaje de escritura en un valor inferior para limitar las operaciones de Amazon EMR. Para obtener información sobre la configuración de rendimiento de DynamoDB, consulte Rendimiento aprovisionado.
Si una tabla está configurada para el modo bajo demanda, debe volver a cambiarlo al modo aprovisionado antes de ejecutar una operación de exportación o importación. Las tuberías necesitan una relación de rendimiento para calcular los recursos que se van a utilizar de un Dynamo. DBtable El modo bajo demanda elimina el rendimiento aprovisionado. Para aprovisionar la capacidad de rendimiento, puede utilizar las métricas de Amazon CloudWatch Events para evaluar el rendimiento agregado que ha utilizado una tabla.

Temas

- Configurar una tabla de Hive para ejecutar comandos de Hive
- Ejemplo de comandos de Hive para la exportación, importación y consulta de datos en DynamoDB
- Optimización de rendimiento para las operaciones de Amazon EMR en DynamoDB

Configurar una tabla de Hive para ejecutar comandos de Hive

Apache Hive es una aplicación de almacenamiento de datos que se puede utilizar para consultar los datos contenidos en es de Amazon EMR con un lenguaje de tipo SQL. Para obtener más información acerca de Hive, consulte http://hive.apache.org.

El siguiente procedimiento supone que ya ha creado un clúster y ha especificado un par de EC2 claves de Amazon. Para obtener información sobre cómo empezar a crear clústeres, consulte Introducción a Amazon EMR en la Guía de administración de Amazon EMR.

Configure Hive para usar MapReduce

Cuando se usa Hive en Amazon EMR para consultar tablas de DynamoDB, se pueden producir errores si Hive usa el motor de ejecución predeterminado, Tez. Por este motivo, cuando cree

un clúster con Hive que se integre con DynamoDB, tal como se describe en esta sección, le recomendamos que utilice una clasificación de configuración que defina el uso de Hive. MapReduce Para obtener más información, consulte Configuración de aplicaciones.

En el siguiente fragmento se muestran la clasificación y la propiedad de la configuración que se van a utilizar para configurarla MapReduce como motor de ejecución de Hive:

```
[
                {
                    "Classification": "hive-site",
                    "Properties": {
                        "hive.execution.engine": "mr"
                    }
                }
            ]
```

Para ejecutar comandos de Hive de forma interactiva

1. Conecte al nodo principal. Para obtener más información, consulte Conexión al nodo maestro mediante SSH en la Guía de administración de Amazon EMR.
2. En el símbolo del sistema de comandos para el nodo principal, escriba `hive`.

   Debería ver un símbolo del sistema de Hive: `hive>`
3. Escriba un comando de Hive que asigna una tabla en la aplicación de Hive a los datos de DynamoDB. Esta tabla actúa como una referencia para los datos almacenados en Amazon DynamoDB; los datos no se almacenan localmente en Hive y las consultas que utilizan esta tabla se ejecutan frente a los datos en vivo en DynamoDB, consumiendo capacidad de lectura o escritura de la tabla cada vez que se ejecuta un comando. Si espera ejecutar varios comandos de Hive respecto al mismo conjunto de datos, plantéese en primer lugar la opción de exportar.

   A continuación, se muestra la sintaxis para el mapeo de una tabla de Hive a una tabla de DynamoDB.

   ```
   CREATE EXTERNAL TABLE hive_tablename
    (hive_column1_name column1_datatype, hive_column2_name column2_datatype...)
   STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
   TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
   "dynamodb.column.mapping" =
    "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name..."
   ```

Al crear una tabla en Hive desde DynamoDB, debe crearla como tabla externa mediante la palabra clave EXTERNAL. La diferencia entre las tablas internas y externas es que los datos en tablas internas se eliminan cuando se anula una tabla interna. Este no es el comportamiento deseado cuando hay conexión a Amazon DynamoDB y, por lo tanto, solo se admiten las tablas externas.

Por ejemplo, el siguiente comando de Hive crea una tabla denominada hivetable1 en Hive que hace referencia a la tabla de DynamoDB denominada dynamodbtable1. La tabla dynamodbtable1 de DynamoDB tiene un esquema de clave principal. hash-and-range El elemento de clave hash es `name` (tipo cadena), el elemento de clave de rango es `year` (tipo numérico) y cada elemento tiene un valor de atributo para `holidays` (tipo conjunto de cadenas).

```
CREATE EXTERNAL TABLE hivetable1 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");
```

La línea 1 utiliza la instrucción `CREATE EXTERNAL TABLE` de HiveQL. Para hivetable1, debe establecer una columna para cada par de atributos nombre-valor en la tabla de DynamoDB y proporcionar el tipo de datos. Estos valores no distinguen entre mayúsculas y minúsculas y puede proporcionar a las columnas cualquier nombre (excepto palabras reservadas).

La línea 2 utiliza la instrucción `STORED BY`. El valor `STORED BY` es el nombre de la clase que se encarga de la conexión entre Hive y DynamoDB. Debe definirse en `'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'`.

La línea 3 utiliza la instrucción `TBLPROPERTIES` para asociar "hivetable1" con la tabla y el esquema correctos en DynamoDB. Proporcione a `TBLPROPERTIES` valores para el parámetro `dynamodb.table.name` y el parámetro `dynamodb.column.mapping`. Estos valores distinguen entre mayúsculas y minúsculas.

> **Note**
>
> Todos los nombres de atributos de DynamoDB de la tabla deben tener las columnas correspondientes en la tabla de Hive. Según la versión de Amazon EMR, se producen las siguientes situaciones si el one-to-one mapeo no existe:
>
> - En la versión 5.27.0 y posteriores de Amazon EMR, el conector tiene validaciones que garantizan un mapeo one-to-one entre los nombres de los atributos de DynamoDB y las columnas de la tabla Hive. Si la asignación no existe, se producirá un error. one-to-one
> - En la versión 5.26.0 de Amazon EMR y anteriores, la tabla de Hive no incluirá el par nombre-valor de DynamoDB. Si no asigna los atributos de clave principal de DynamoDB, Hive devuelve un error. Si no asigna un atributo de clave no principal, no se genera ningún error, pero no verá los datos en la tabla de Hive. Si los tipos de datos no coinciden, el valor es nulo.

A continuación, puede comenzar a ejecutar operaciones de Hive en hivetable1. Las consultas que se ejecutan en hivetable1 se ejecutan internamente en la tabla de DynamoDB dynamodbtable1 de su cuenta de DynamoDB, y consumen unidades de lectura o escritura con cada ejecución.

Cuando ejecuta consultas de Hive frente a una tabla de DynamoDB, debe asegurarse de que ha aprovisionado una cantidad suficiente de unidades de capacidad de lectura.

Por ejemplo, suponga que ha aprovisionado 100 unidades de capacidad de lectura para su tabla de DynamoDB. Esto le permite realizar 100 lecturas o 409 600 bytes, por segundo. Si dicha tabla contiene 20 GB de datos (21 474 836 480 bytes) y la consulta de Hive realiza un análisis completo de la tabla, puede calcular aproximadamente cuánto tardará la consulta en ejecutarse:

21 474 836 480 / 409 600 = 52 429 segundos = 14,56 horas

La única forma de reducir el tiempo necesario, sería ajustar las unidades de capacidad de lectura en el origen de la tabla de DynamoDB. Agregar más nodos de Amazon EMR no le ayudará.

En la salida de Hive, el porcentaje completado se actualiza cuando finalizan uno o varios procesos de mapeador. Para una tabla de DynamoDB grande con un ajuste bajo de capacidad de lectura aprovisionada, la salida de porcentaje de finalización podría no actualizarse durante mucho tiempo; en el caso anterior, el trabajo aparecerá con un 0% completo durante varias horas. Para ver un

estado más detallado del progreso del trabajo, acceda a la consola de Amazon EMR; podrá ver el estado de la tarea de asignador individual y estadísticas de lecturas de datos. También puede iniciar sesión en la interfaz de usuario de Hadoop en el nodo principal y ver las estadísticas de Hadoop. Esto le mostrará el estado de tareas de mapeo individuales y algunos datos de estadísticas de lectura. Para obtener más información, consulte los temas siguientes:

- Interfaces web alojadas en el nodo maestro
- Ver las interfaces web de Hadoop

Para obtener más información sobre las instrucciones de HiveQL de ejemplo para realizar tareas tales como la exportación o importación de datos desde DynamoDB y la unión de tablas, consulte Ejemplo de comandos de Hive para la exportación, importación y consulta de datos en DynamoDB.

Para cancelar una solicitud de Hive

Cuando se ejecuta una consulta de Hive, la respuesta inicial del servidor incluye el comando para cancelar la solicitud. Para cancelar la solicitud en cualquier momento del proceso, utilice Kill Command (Comando de cancelación) de la respuesta del servidor.

1. Introduzca `Ctrl+C` para salir del cliente de línea de comando.
2. En el símbolo del shell, introduzca el Kill Command (Comando de cancelación) de la respuesta inicial del servidor a su solicitud.

   Como alternativa, puede ejecutar el siguiente comando desde la línea de comandos del nodo maestro para eliminar el trabajo de Hadoop, que *job-id* es el identificador del trabajo de Hadoop y se puede recuperar desde la interfaz de usuario de Hadoop.

   ```
   hadoop job -kill job-id
   ```

Tipos de datos para Hive y DynamoDB

En la siguiente tabla se muestran los tipos de datos de Hive disponibles, el tipo predeterminado de DynamoDB al que se corresponden y los tipos alternativos de DynamoDB a los que también se pueden asignar.

Tipo de Hive	Tipo predeterminado de DynamoDB	Tipos alternativos de DynamoDB
cadena	cadena (S)	
bigint o doble	número (N)	
binario	binario (B)	
booleano	booleano (BOOL)	
array	lista (L)	conjunto de números (NS), un conjunto de cadenas (SS) o un conjunto de binarios (BS)
mapa <string,string>	elemento	mapa (M)
mapa <string,?>	mapa (M)	
	nulo (NULL)	

Si desea escribir los datos de Hive como un correspondiente tipo alternativo de DynamoDB o si los datos de DynamoDB incluyen valores de atributo de un tipo alternativo de DynamoDB, puede especificar la columna y el tipo de DynamoDB con el parámetro `dynamodb.type.mapping`. En el siguiente ejemplo se muestra la sintaxis para especificar un mapeo de tipo alternativo.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.type.mapping" = "hive_column1_name:dynamodb_attribute1_datatype");
```

El parámetro de mapeo de tipos es opcional y solo tiene que especificarse para las columnas que utilizan tipos alternativos.

Por ejemplo, el siguiente comando de Hive crea una tabla denominada `hivetable2` que hace referencia a la tabla `dynamodbtable2` de DynamoDB. Es similar a `hivetable1`, salvo que asigna la columna `col3` al tipo de conjunto de cadenas (SS).

```
CREATE EXTERNAL TABLE hivetable2 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable2",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays",
"dynamodb.type.mapping" = "col3:SS");
```

En Hive, `hivetable1` y `hivetable2` son idénticos. Sin embargo, cuando los datos de esas tablas se escriben en sus correspondientes tablas de DynamoDB, `dynamodbtable1` incluirá listas, mientras que `dynamodbtable2` incluirá conjuntos de cadenas.

Si desea escribir valores `null` de Hive como atributos de tipo `null` de DynamoDB, puede hacerlo con el parámetro `dynamodb.null.serialization`. En el siguiente ejemplo se muestra la sintaxis para especificar la serialización `null`.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.null.serialization" = "true");
```

El parámetro de serialización null es opcional y se establece en `false` si no se especifica. Tenga en cuenta que los atributos `null` de DynamoDB se leen como valores `null` en Hive con independencia de la configuración del parámetro. Las colecciones de Hive con valores `null` solo pueden escribirse en DynamoDB si el parámetro de serialización null se especifica como `true`. De lo contrario, se produce un error de Hive.

El tipo bigint en Hive es igual al tipo long en Java y el tipo de doble de Hive es igual que el tipo doble de Java en términos de precisión. Esto significa que, si dispone de datos numéricos almacenados en DynamoDB con precisión superior a la disponible en los tipos de datos de Hive, el uso de Hive para exportar, importar o hacer referencia a los datos de DynamoDB podría dar lugar a una pérdida de precisión o un error en la consulta de Hive.

Las exportaciones del tipo binario desde DynamoDB a Amazon Simple Storage Service (Amazon S3) o HDFS se almacenan como cadena codificada en Base64. Si está importando datos desde Amazon S3 o HDFS al tipo binario de DynamoDB, deberían estar codificados como una cadena Base64.

Opciones de Hive

Puede definir las siguientes opciones de Hive para gestionar la transferencia de datos fuera de Amazon DynamoDB. Estas opciones solo persisten para la sesión de Hive actual. Si cierra el símbolo del sistema de Hive y vuelve a abrirlo más adelante en el clúster, estos ajustes habrán vuelto a los valores predeterminados.

Opciones de Hive	Descripción
`dynamodb.throughput.read.percent`	Define la velocidad de las operaciones de lectura para mantener la tasa de rendimiento aprovisionada de DynamoDB en el rango asignado para la tabla. El valor está entre `0.1` y `1.5`, ambos incluidos. El valor de 0,5 es la tasa de lectura predeterminada, lo que significa que Hive intentará consumir la mitad de la lectura aprovisionada a través de los recursos de la tabla. Al aumentar este valor por encima de 0,5 se aumenta la tasa de solicitud de lectura. Si se reduce el valor por debajo de 0,5 se reduce la tasa de solicitud de lectura. Esta tasa de lectura es aproximada. La tasa de lectura real dependerá de factores tales como el hecho de que exista una distribución uniforme de claves en DynamoDB. Si encuentra que la operación de Hive suele superar el rendimiento aprovisionado o si el tráfico de lectura en directo se está demasiado limitado, entonces reduzca este valor por debajo de `0.5`. Si dispone de suficiente capacidad y desea una operación de Hive más rápida, defina este valor por encima de `0.5`. También puede suscribir en exceso definiendo este valor en 1,5 si cree que hay operaciones de entrada/salida sin utilizar disponibles.

Opciones de Hive	Descripción
`dynamodb.throughput.write.percent`	Establece la velocidad de las operaciones de escritura para mantener la tasa de rendimiento aprovisionada de DynamoDB en el rango asignado para la tabla. El valor está entre `0.1` y `1.5`, ambos incluidos. El valor de 0,5 es la tasa de escritura predeterminada, lo que significa que Hive intentará consumir la mitad de la escritura aprovisionada a través de los recursos de la tabla. Al aumentar este valor por encima de 0,5 se aumenta la tasa de solicitud de escritura. Si se reduce el valor por debajo de 0,5 se reduce la tasa de solicitud de escritura. Esta tasa de escritura es aproximada. La tasa de escritura real dependerá de factores tales como el hecho de que exista una distribución uniforme de claves en DynamoDB Si encuentra que la operación de Hive suele superar el rendimiento aprovisionado o si el tráfico de escritura en directo se está demasiado limitado, entonces reduzca este valor por debajo de `0.5`. Si dispone de suficiente capacidad y desea una operación de Hive más rápida, defina este valor por encima de `0.5`. También puede suscribir en exceso definiendo este valor en 1,5 si cree que hay operaciones de entrada/salida sin utilizar disponibles o que estos son los datos iniciales cargados en la tabla y no hay tráfico activo todavía.
`dynamodb.endpoint`	Especifique el punto de conexión del servicio de DynamoDB. Para obtener más información sobre los puntos de conexión de DynamoDB disponibles, consulte Regiones y puntos de conexión.

Opciones de Hive	Descripción
`dynamodb.max.map.tasks`	Especifique el número máximo de tareas de asignación al leer datos desde DynamoDB. Este valor debe ser igual o superior a 1.
`dynamodb.retry.duration`	Especifique el número de minutos que utilizar como duración de tiempo de inactividad para volver a intentar ejecutar los comandos de Hive. Este valor debe ser un entero igual o superior a 0. El tiempo de espera predeterminado tiene una duración de dos minutos.

Estas opciones se establecen utilizando el comando `SET` tal y como se muestra en el siguiente ejemplo.

```
SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Ejemplo de comandos de Hive para la exportación, importación y consulta de datos en DynamoDB

Los siguientes ejemplos utilizan comandos de Hive para llevar a cabo operaciones como: exportación de datos a Amazon S3 o HDFS, importación de datos a DynamoDB, unión de tablas, consulta de tablas y mucho más.

Las operaciones en una tabla de Hive hacen referencia los datos almacenados en DynamoDB. Los comandos de Hive están sujetos a la configuración de rendimiento aprovisionado de la tabla de DynamoDB y los datos recuperados incluyen los datos escritos en la tabla de DynamoDB en el momento en que DynamoDB procesa la solicitud de la operación de Hive. Si el proceso de recuperación de datos tarda mucho tiempo, algunos datos devueltos por el comando Hive es posible que se hayan actualizado en DynamoDB desde que se inició el comando de Hive.

Los comandos de Hive `DROP TABLE` y `CREATE TABLE` solo actúan sobre la tablas locales en Hive y no crean ni borran tablas en DynamoDB. Si su consulta de Hive hace referencia a una tabla en DynamoDB, la tabla ya existe antes de ejecutar la consulta. Para obtener más información sobre la creación y eliminación de tablas en DynamoDB, consulte, Uso de tablas en DynamoDB en la Guía para desarrolladores de Amazon DynamoDB.

> **Note**
>
> Al mapear una tabla de Hive a una ubicación de Amazon S3, no la mapee a la ruta raíz del bucket, s3://amzn-s3-demo-bucket, ya que esto puede provocar errores cuando Hive escriba los datos en Amazon S3. En lugar de ello asigne la tabla a una subruta del bucket, s3://amzn-s3-demo-bucket/mypath.

Exportación de datos desde DynamoDB

Puede utilizar Hive para exportar datos desde DynamoDB.

Para exportar una tabla de DynamoDB a un bucket de Amazon S3

- Cree una tabla de Hive que haga referencia a los datos almacenados en DynamoDB. A continuación, puede llamar al comando INSERT OVERWRITE para escribir los datos en un directorio externo. En el siguiente ejemplo, *s3://amzn-s3-demo-bucket/path/subpath/* es una ruta válida en Amazon S3. Ajuste las columnas y tipos de datos en el comando CREATE para asignar los valores en su DynamoDB. Puede utilizar esto para crear un archivo de los datos de DynamoDB en Amazon S3.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


INSERT OVERWRITE DIRECTORY 's3://amzn-s3-demo-bucket/path/subpath/' SELECT *
FROM hiveTableName;
```

Para exportar una tabla de DynamoDB a un bucket de Amazon S3 utilizando del formato

- Cree una tabla externa que haga referencia a una ubicación en Amazon S3. Esto se muestra a continuación como s3_export. Durante la llamada CREATE, especifique el formato de fila para la tabla. A continuación, cuando utilice INSERT OVERWRITE para exportar datos desde DynamoDB a s3_export, los datos se escriben en el formato especificado. En el siguiente ejemplo, los datos se escriben como valores separados por comas (CSV).

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Para exportar una tabla de DynamoDB a un bucket de Amazon S3 sin especificar un mapeo de columnas

- Cree una tabla de Hive que haga referencia a los datos almacenados en DynamoDB. Esto es similar al ejemplo anterior, salvo que no tiene que especificar un mapeo de columnas. En la tabla debe tener exactamente una columna de tipo `map<string, string>`. Si crea una tabla EXTERNAL en Amazon S3, puede llamar al comando `INSERT OVERWRITE` para escribir los datos de DynamoDB en Amazon S3. Puede utilizar esto para crear un archivo de los datos de DynamoDB en Amazon S3. Dado que no hay ningún mapeo de columnas, no puede consultar las tablas que se hayan exportado de esta manera. La exportación de datos sin especificar un mapeo de columna está disponible en Hive 0.8.1.5 o versiones posteriores, que es compatible con la AMI 2.2.x y posteriores de Amazon EMR.

```
CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
```

```
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3TableName SELECT *
FROM hiveTableName;
```

Para exportar una tabla de DynamoDB a un bucket de Amazon S3 utilizando la compresión de datos

- Hive proporciona varios códecs de compresión que puede definir durante la sesión de Hive. Al hacerlo los datos exportados se comprimen en el formato especificado. En el siguiente ejemplo, se comprimen los archivos exportados mediante el Lempel-Ziv-Oberhumer algoritmo (LZO).

```
SET hive.exec.compress.output=true;
SET io.seqfile.compression.type=BLOCK;
SET mapred.output.compression.codec = com.hadoop.compression.lzo.LzopCodec;


CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE lzo_compression_table (line STRING)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE lzo_compression_table SELECT *
FROM hiveTableName;
```

  Los códecs de compresión disponibles son:

  - org.apache.hadoop.io.compress. GzipCodec
  - org.apache.hadoop.io.compress. DefaultCodec

- com.hadoop.compression.lzo. LzoCodec
- com.hadoop.compression.lzo. LzopCodec
- org.apache.hadoop.io.compress. BZip2Códec
- org.apache.hadoop.io.compress. SnappyCodec

Para exportar una tabla de DynamoDB a HDFS

- Utilice el siguiente comando Hive, donde *hdfs:///directoryName* hay una ruta HDFS válida y *hiveTableName* es una tabla de Hive que hace referencia a DynamoDB. Esta operación de exportación es más rápida que la exportación de una tabla de DynamoDB a Amazon S3 debido a que Hive 0.7.1.1 utiliza HDFS como paso intermedio al exportar datos a Amazon S3. El siguiente ejemplo muestra también cómo definir `dynamodb.throughput.read.percent` en 1,0 para aumentar la tasa de solicitud de lectura.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE DIRECTORY 'hdfs:///directoryName' SELECT * FROM hiveTableName;
```

  También puede exportar datos a HDFS mediante el formato y la compresión que se muestran más arriba para la exportación a Amazon S3. Para hacerlo, solo tiene que sustituir el directorio de Amazon S3 en los ejemplos anteriores por un directorio HDFS.

Para leer datos de caracteres UTF-8 no imprimibles en Hive

- Puede leer y escribir datos de caracteres UTF-8 no imprimibles con Hive utilizando la cláusula `STORED AS SEQUENCEFILE` al crear la tabla. A SequenceFile es el formato de archivo binario de Hadoop; debe usar Hadoop para leer este archivo. En el siguiente ejemplo se muestra cómo exportar datos desde DynamoDB a Amazon S3. Puede utilizar esta funcionalidad para controlar caracteres codificados con UTF-8 no imprimibles.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
STORED AS SEQUENCEFILE
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

Importación de datos a DynamoDB

Al escribir datos en DynamoDB utilizando Hive, debe asegurarse de que el número de unidades de capacidad de escritura sea mayor que el número de asignadores en el clúster. Por ejemplo, los clústeres que se ejecutan en instancias m1.xlarge EC2 producen 8 mapeadores por instancia. En el caso de un clúster; que tiene 10 instancias, eso supondría un total de 80 asignadores. Si sus unidades de capacidad de escritura no son mayores que el número de asignadores en el clúster, la operación de escritura de Hive puede consumir todo el rendimiento de escritura o intentar consumir más rendimiento del aprovisionado. Para obtener más información sobre la cantidad de mapeadores que produce cada EC2 tipo de instancia, consulte. Configurar Hadoop

El número de mapeadores de Hadoop se controla mediante las divisiones de entrada. Si hay muy pocas divisiones, el comando de escritura podría no ser capaz de consumir todo el rendimiento de escritura disponible.

Si existe un elemento con la misma clave en la tabla de DynamoDB de destino, se sobrescribe. Si no existe ningún elemento con la clave en la tabla de DynamoDB de destino, el elemento se inserta.

Para importar datos de Amazon S3 a DynamoDB

- Puede usar Amazon EMR (Amazon EMR) y Hive para escribir datos de Amazon S3 en DynamoDB.

```
CREATE EXTERNAL TABLE s3_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM s3_import;
```

Para importar una tabla desde un bucket de Amazon S3 a DynamoDB sin especificar un mapeo de columnas

- Cree una tabla EXTERNAL que haga referencia a los datos almacenados en Amazon S3 que se exportaron previamente desde DynamoDB. Antes de la importación, asegúrese de que la tabla existe en DynamoDB y que tiene el mismo esquema de claves que la tabla de DynamoDB exportada anteriormente. Además, la tabla debe tener exactamente una columna de tipo `map<string, string>`. Si, a continuación, crea una tabla de Hive que está vinculada a DynamoDB, puede llamar al comando `INSERT OVERWRITE` para escribir los datos de Amazon S3 en DynamoDB. Dado que no hay ningún mapeo de columnas, no puede consultar las tablas que se hayan importado de esta manera. La importación de datos sin especificar un mapeo de columna está disponible en Hive 0.8.1.5 o versiones posteriores, que es compatible con la AMI 2.2.3 y posteriores de Amazon EMR.

```
CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

INSERT OVERWRITE TABLE hiveTableName SELECT *
FROM s3TableName;
```

Para importar una tabla desde HDFS a DynamoDB

- Puede utilizar Amazon EMR y Hive para escribir datos de HDFS en DynamoDB.

```
CREATE EXTERNAL TABLE hdfs_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 'hdfs:///directoryName';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM hdfs_import;
```

Consulta de datos en DynamoDB

Los siguientes ejemplos muestran las distintas formas en las que puede utilizar Amazon EMR para consultar los datos almacenados en DynamoDB.

Para encontrar el mayor valor para una columna mapeada (**`max`**)

- Utilice los comandos de Hive como los siguientes. En el primer comando, la instrucción CREATE crea una tabla de Hive que hace referencia a los datos almacenados en DynamoDB. La instrucción SELECT utiliza a continuación dicha tabla para consultar los datos almacenados en DynamoDB. El siguiente ejemplo encuentra el mayor pedido realizado por un determinado cliente.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT max(total_cost) from hive_purchases where customerId = 717;
```

Para agregar datos utilizando la cláusula **GROUP BY**

- Puede utilizar la cláusula GROUP BY para recopilar datos a través de varios registros. Esto se utiliza a menudo con una función de agregación, como suma, recuento, mín o máx. En el siguiente ejemplo, se devuelve una lista de los mayores pedidos de clientes que han realizado más de tres pedidos.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT customerId, max(total_cost) from hive_purchases GROUP BY customerId HAVING
 count(*) > 3;
```

Para unir dos tablas de DynamoDB

- El siguiente ejemplo asocia dos tablas de Hive a los datos almacenados en DynamoDB. A continuación, hace una llamada para unir estas dos tablas. La unión se calcula en el clúster y se devuelve. La unión no tiene lugar en DynamoDB. Este ejemplo devuelve una lista de los clientes y de sus compras para clientes que han realizado más de dos pedidos.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

CREATE EXTERNAL TABLE hive_customers(customerId bigint, customerName string,
 customerAddress array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Customers",
"dynamodb.column.mapping" =
 "customerId:CustomerId,customerName:Name,customerAddress:Address");
```

```
Select c.customerId, c.customerName, count(*) as count from hive_customers c
JOIN hive_purchases p ON c.customerId=p.customerId
GROUP BY c.customerId, c.customerName HAVING count > 2;
```

Para unir dos tablas desde orígenes diferentes

- En el siguiente ejemplo, Customer_S3 es una tabla de Hive que carga un archivo CSV almacenado en Amazon S3 y hive_purchases es una tabla que hace referencia a datos de DynamoDB. El siguiente ejemplo une datos de clientes almacenados en un archivo CSV en Amazon S3 con datos de pedidos almacenados en DynamoDB para devolver un conjunto de datos que representa los pedidos realizados por los clientes que tienen "Miller" en su nombre.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

CREATE EXTERNAL TABLE Customer_S3(customerId bigint, customerName string,
 customerAddress array<String>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

Select c.customerId, c.customerName, c.customerAddress from
Customer_S3 c
JOIN hive_purchases p
ON c.customerid=p.customerid
where c.customerName like '%Miller%';
```

> **Note**
>
> En los ejemplos anteriores, se incluyeron las instrucciones CREATE TABLE en cada ejemplo para mayor claridad e integridad. Al ejecutar varias consultas o exportar operaciones en una

tabla de Hive determinada, solo tiene que crear la tabla una vez, al principio de la sesión de Hive.

Optimización de rendimiento para las operaciones de Amazon EMR en DynamoDB

Las operaciones de Amazon EMR en una tabla de DynamoDB cuentan como operaciones de lectura y están sujetas a la configuración de rendimiento aprovisionado de la tabla. Amazon EMR implementa su propia lógica para intentar balancear la carga en su tabla de DynamoDB para minimizar la posibilidad de sobrepasar el rendimiento aprovisionado. Al final de cada consulta de Hive, Amazon EMR devuelve información sobre el clúster; utilizado para procesar la consulta, incluyendo el número de veces que se ha superado el rendimiento aprovisionado. Puede utilizar esta información, así como CloudWatch las métricas sobre el rendimiento de DynamoDB, para gestionar mejor la carga de la tabla de DynamoDB en solicitudes posteriores.

Los siguientes factores influyen en el rendimiento de las consultas de Hive cuando se trabaja con tablas de DynamoDB.

Unidades de capacidad de lectura aprovisionada

Cuando ejecuta consultas de Hive frente a una tabla de DynamoDB, debe asegurarse de que ha aprovisionado una cantidad suficiente de unidades de capacidad de lectura.

Por ejemplo, suponga que ha aprovisionado 100 unidades de capacidad de lectura para su tabla de DynamoDB. Esto le permite realizar 100 lecturas o 409 600 bytes, por segundo. Si dicha tabla contiene 20 GB de datos (21 474 836 480 bytes) y la consulta de Hive realiza un análisis completo de la tabla, puede calcular aproximadamente cuánto tardará la consulta en ejecutarse:

21 474 836 480 / 409 600 = 52 429 segundos = 14,56 horas

La única forma de reducir el tiempo necesario, sería ajustar las unidades de capacidad de lectura en el origen de la tabla de DynamoDB. Agregar más nodos al clúster de Amazon EMR no servirá.

En la salida de Hive, el porcentaje completado se actualiza cuando finalizan uno o varios procesos de mapeador. Para una tabla de DynamoDB grande con un ajuste bajo de capacidad de lectura aprovisionada, la salida de porcentaje de finalización podría no actualizarse durante mucho tiempo; en el caso anterior, el trabajo aparecerá con un 0% completo durante varias horas. Para ver un

estado más detallado del progreso del trabajo, acceda a la consola de Amazon EMR; podrá ver el estado de la tarea de asignador individual y estadísticas de lecturas de datos.

También puede iniciar sesión en la interfaz de usuario de Hadoop en el nodo principal y ver las estadísticas de Hadoop. Esto le muestra el estado de tareas de mapeo individuales y algunos datos de estadísticas de lectura. Para obtener más información, consulte Interfaces web alojadas en el nodo maestro en la Guía de administración de Amazon EMR.

Ajuste de porcentaje de lectura

De forma predeterminada, Amazon EMR administra la carga de solicitudes de la tabla de DynamoDB de acuerdo con el rendimiento aprovisionado actual. Sin embargo, cuando Amazon EMR devuelve información sobre su trabajo que incluye un elevado número de respuestas que superan el rendimiento aprovisionado, puede ajustar la tasa de lectura predeterminada mediante el parámetro `dynamodb.throughput.read.percent` cuando configure la tabla de Hive. Para obtener más información acerca cómo configurar el parámetro de porcentaje de lectura, consulte Opciones de Hive.

Configuración de porcentaje de escritura

De forma predeterminada, Amazon EMR administra la carga de solicitudes de la tabla de DynamoDB de acuerdo con el rendimiento aprovisionado actual. Sin embargo, cuando Amazon EMR devuelve información sobre su trabajo que incluye un elevado número de respuestas que superan el rendimiento aprovisionado, puede ajustar la tasa de escritura predeterminada mediante el parámetro `dynamodb.throughput.write.percent` cuando configure la tabla de Hive. Para obtener más información acerca cómo configurar el parámetro de porcentaje de escritura, consulte Opciones de Hive.

Configuración de duración de reintento

De forma predeterminada, Amazon EMR vuelve a ejecutar una consulta de Hive si no devuelve un resultado en el plazo de dos minutos, el intervalo de reintento predeterminado. Puede ajustar este intervalo configurando el parámetro `dynamodb.retry.duration` al ejecutar una consulta de Hive. Para obtener más información acerca cómo configurar el parámetro de porcentaje de escritura, consulte Opciones de Hive.

Número de tareas de asignación

Los daemons de asignador que lanza Hadoop para procesar las solicitudes para exportar y consultar datos almacenados en DynamoDB están limitados a una tasa de lectura máxima de 1 MiB por

segundo para limitar la capacidad de lectura utilizada. Si dispone de rendimiento aprovisionado adicional en DynamoDB, puede mejorar el rendimiento de las operaciones de consulta y exportación de Hive aumentando el número de daemons de asignador. Para ello, puede aumentar el número de EC2 instancias del clúster o aumentar el número de daemons mapeadores que se ejecutan en cada instancia. EC2

Para aumentar el número de EC2 instancias de un clúster, debes detener el clúster actual y volver a lanzarlo con un número mayor de instancias. EC2 El número de EC2 instancias se especifica en el cuadro de diálogo Configurar EC2 instancias si se lanza el clúster desde la consola de Amazon EMR, o con la `--num-instances` opción si se lanza el clúster desde la CLI.

La cantidad de tareas de mapa que se ejecutan en una instancia depende del tipo de EC2 instancia. Para obtener más información sobre los tipos de EC2 instancias compatibles y la cantidad de mapeadores que proporciona cada uno, consulteConfiguración de tareas. Allí encontrará una sección "Task Configuration" para cada una de las configuraciones admitidas.

Otra forma de aumentar el número de daemons de mapeador consiste en cambiar el parámetro de configuración `mapreduce.tasktracker.map.tasks.maximum` de Hadoop por un valor mayor. Esto tiene la ventaja de ofrecer más mapeadores sin aumentar el número ni el tamaño de las EC2 instancias, lo que le permite ahorrar dinero. Una desventaja es que establecer este valor demasiado alto puede provocar que las EC2 instancias del clúster se queden sin memoria. Para configurar `mapreduce.tasktracker.map.tasks.maximum`, lance el clúster y especifique un valor para `mapreduce.tasktracker.map.tasks.maximum` como una propiedad de la clasificación de configuración mapred-site. Esto se muestra en el siguiente ejemplo. Para obtener más información, consulte Configuración de aplicaciones.

```
{
    "configurations": [
    {
        "classification": "mapred-site",
        "properties": {
            "mapred.tasktracker.map.tasks.maximum": "10"
        }
    }
    ]
}
```

Solicitudes de datos paralelas

Varias solicitudes de datos, ya sea de más de un usuario o más de una aplicación en una única tabla podría agotar el rendimiento de lectura aprovisionado y un dar lugar a un rendimiento lento.

Duración del proceso

La consistencia de datos en DynamoDB depende del orden de las operaciones de lectura y escritura en cada nodo. Aunque haya una consulta de Hive en curso, otra aplicación podría cargar nuevos datos en la tabla de DynamoDB o modificar o eliminar datos existentes. En este caso, los resultados de la consulta de Hive podrían no reflejar los cambios realizados en los datos mientras se ejecutaba la consulta.

Evitar un rendimiento excesivo

Al ejecutar consultas de Hive en DynamoDB, tenga cuidado de no superar el rendimiento aprovisionado, ya que esto agotará la capacidad necesaria para las llamadas de la aplicación a `DynamoDB::Get`. Para asegurarte de que esto no ocurra, deberías controlar periódicamente el volumen de lectura y la limitación de las llamadas a las aplicaciones `DynamoDB::Get` consultando los registros y monitorizando las métricas en Amazon. CloudWatch

Tiempo de solicitud

Programar las consultas de Hive que acceden a una tabla de DynamoDB de modo que se lleven a cabo en aquellos momentos en que la demanda de esa tabla de DynamoDB es más baja, mejora el rendimiento. Por ejemplo, si la mayoría de los usuarios de la aplicación viven en San Francisco, podría elegir exportar los datos cada día a las 4:00 h PST, cuando la mayoría de los usuarios están durmiendo, y no actualizan los registros de la base de datos de DynamoDB.

Tablas basadas en tiempo

Si los datos se organizan como una serie de tablas de DynamoDB basadas en tiempo como, por ejemplo, una tabla al día, puede exportar los datos cuando la tabla deje de estar activa. Puede utilizar esta técnica para crear copias de seguridad de los datos en Amazon S3 de manera continuada.

Datos archivados

Si tiene previsto ejecutar muchas consultas de Hive en los datos almacenados en DynamoDB y la aplicación puede tolerar los datos archivados, es posible que desee exportar los datos a HDFS o a

Amazon S3 y ejecutar las consultas de Hive en una copia de los datos en lugar de DynamoDB. Esto conserva las operaciones de lectura y el rendimiento aprovisionado.

Kinesis

Los clústeres de Amazon EMR pueden leer y procesar las transmisiones de Amazon Kinesis directamente mediante herramientas conocidas del ecosistema de Hadoop, como Hive, Pig MapReduce, la API de streaming de Hadoop y Cascading. También puede unir los datos en tiempo real de Amazon Kinesis con los datos existentes en Amazon S3, Amazon DynamoDB y HDFS en un clúster en ejecución. Puede cargar datos directamente de Amazon EMR en Amazon S3 o en DynamoDB para actividades posteriores al procesamiento. Para obtener información sobre las características y los precios del servicio Amazon Kinesis, consulte la página de Amazon Kinesis.

¿Qué puede hacer con la integración de Amazon EMR y Amazon Kinesis?

La integración entre Amazon EMR y Amazon Kinesis facilita en gran medida ciertas actividades, como, por ejemplo:

- Análisis de registros de streaming: puede analizar registros web de streaming para generar una lista de los 10 tipos de error principales cada pocos minutos por región, navegador y dominio de acceso.
- Interacción con los clientes: puede escribir consultas que unan datos de secuencias de clic desde Amazon Kinesis con información de campañas publicitarias almacenada en una tabla de DynamoDB para identificar las categorías de anuncios más eficaces que se muestran en sitios web particulares.
- Consultas interactivas ad-hoc: puede cargar datos periódicamente desde flujos de Amazon Kinesis en HDFS y hacer que estén disponibles como tabla de Impala local para consultas analíticas interactivas rápidas.

Análisis con puntos de control de flujos de Amazon Kinesis

Los usuarios pueden ejecutar análisis por lotes periódicos de flujos de Amazon Kinesis en lo que se denominan iteraciones. Dado que los registros de datos de los flujos de Amazon Kinesis se recuperan utilizando un número de secuencia, los límites de las iteraciones se definen por números de secuencia iniciales y finales que Amazon EMR almacena en una tabla de DynamoDB. Por ejemplo, cuando finaliza `iteration0`, almacena el número de secuencia final en DynamoDB para

que cuando comience el trabajo `iteration1`, este pueda recuperar los datos siguientes del flujo. Este mapeo de iteraciones en los datos de flujo se denomina creación de puntos de comprobación. Para obtener más información, consulte Conector de Kinesis.

Si una iteración ha sido objeto de control y el trabajo no ha podido procesar una iteración, Amazon EMR intenta volver a procesar los registros de esa iteración.

La creación de puntos de comprobación es una característica que le permite:

- Comenzar el procesamiento de datos después de un número de secuencia procesado por una consulta anterior que se ejecutó en el mismo flujo y nombre lógico
- Volver a procesar el mismo lote de datos desde Kinesis que se procesó en una consulta anterior

Para habilitar la creación de puntos de comprobación, defina el parámetro `kinesis.checkpoint.enabled` en `true` en sus scripts. Además, configure los siguientes parámetros:

Opción de configuración	Descripción
kinesis.checkpoint.metastore.table.name	Nombre de la tabla de DynamoDB donde se almacenará la información de los puntos de comprobación
kinesis.checkpoint.metastore.hash.key.name	Nombre de clave hash para la tabla de DynamoDB
kinesis.checkpoint.metastore.hash.range.name	Nombre de clave de rango para la tabla de DynamoDB
kinesis.checkpoint.logical.name	Un nombre lógico para el procesamiento actual
kinesis.checkpoint.iteration.no	Número de iteración para procesamiento asociado con el nombre lógico
kinesis.rerun.iteration.without.wait	Valor booleano que indica si una iteración que ha devuelto un error se puede volver a ejecutar sin necesidad de esperar un tiempo de espera; el valor predeterminado es `false`

Recomendaciones de IOPS aprovisionadas para tablas de Amazon DynamoDB

El conector de Amazon EMR para Amazon Kinesis utiliza la base de datos de DynamoDB como respaldo para los metadatos de puntos de comprobación. Debe crear una tabla en DynamoDB antes de consumir los datos de un flujo de Amazon Kinesis con un clúster de Amazon EMR en intervalos de puntos de comprobación. La tabla debe estar en la misma región que su clúster de Amazon EMR. A continuación, se ofrecen recomendaciones generales para el número de IOPS que debe aprovisionar para sus tablas de DynamoDB; `j` es el número máximo de trabajos de Hadoop (con diferente combinación de nombre lógico y número de iteración) que pueden ejecutarse de forma simultánea y `s` es el número máximo de particiones que procesará cualquier trabajo:

Para Read Capacity Units (Unidades de capacidad de lectura): `j*s/5`

Para Write Capacity Units (Unidades de capacidad de escritura): `j*s`

Consideraciones sobre el rendimiento

El rendimiento de particiones de Amazon Kinesis es directamente proporcional al tamaño de instancia de los nodos en los clústeres Amazon EMR y al tamaño de los registros en el flujo. Le recomendamos que use instancias m5.xlarge o superiores en los nodos maestros y principales.

Programación del análisis de Amazon Kinesis con Amazon EMR

Cuando analiza datos de un flujo Amazon Kinesis activo, limitado por los tiempos de espera y una duración máxima de las interacciones, es importante que ejecute el análisis con frecuencia para recopilar detalles del flujo con periodicidad. Hay varias formas de ejecutar estos scripts y consultas en intervalos periódicos. Para tareas recurrentes como estas, recomendamos usar AWS Data Pipeline . Para obtener más información, consulte y en la Guía para desarrolladores. AWS Data Pipeline PigActivityAWS Data Pipeline HiveActivityAWS Data Pipeline

Migración del conector de Spark Kinesis al SDK 2.x para Amazon EMR 7.0

El AWS SDK proporciona un amplio conjunto de APIs bibliotecas para interactuar con los servicios de computación AWS en la nube, como la administración de credenciales y la conexión a los servicios de S3 y Kinesis. El conector de Spark Kinesis se utiliza para consumir datos de Kinesis Data Streams, y los datos recibidos se transforman y procesan en el motor de ejecución de Spark. Actualmente, este conector se basa en la versión 1.x del AWS SDK y Kinesis-client-library (KCL).

Como parte de la migración del AWS SDK 2.x, el conector de Spark Kinesis también se actualizó en consecuencia para que funcione con el SDK 2.x. En la versión 7.0 de Amazon EMR, Spark contiene

la actualización 2.x del SDK que aún no está disponible en la versión comunitaria de Apache Spark. Si utiliza el conector de Spark Kinesis de una versión anterior a la 7.0, debe migrar los códigos de las aplicaciones para que se ejecuten en el SDK 2.x antes de poder migrar a Amazon EMR 7.0.

Guías de migración

En esta sección, se describen los pasos para migrar una aplicación al conector de Spark Kinesis actualizado. Incluye guías para migrar a la biblioteca de clientes de Kinesis (KCL) 2.x, proveedores de AWS credenciales y clientes de AWS servicios en el SDK 2.x. AWS Como referencia, también incluye un WordCountprograma de ejemplo que utiliza el conector Kinesis.

Temas

- Migración de KCL de la versión 1.x a la 2.x
- Migración de proveedores de AWS credenciales del SDK 1.x al AWS 2.x
- Migración de clientes de AWS servicios del AWS SDK 1.x al 2.x
- Ejemplos de código para aplicaciones de transmisión
- Consideraciones a la hora de utilizar el conector de Spark Kinesis actualizado

Migración de KCL de la versión 1.x a la 2.x

- Nivel de métricas y dimensiones en **`KinesisInputDStream`**

  Al crear una instancia de `KinesisInputDStream`, puede controlar el nivel de métricas y las dimensiones para la secuencia. En el siguiente ejemplo, se muestra cómo se pueden personalizar estos parámetros con KCL 1.x:

```
import
 com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel

val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
```

```
  .metricsLevel(MetricsLevel.DETAILED)

 .metricsEnabledDimensions(KinesisClientLibConfiguration.DEFAULT_METRICS_ENABLED_DIMENSIONS.a
  .build()
```

En KCL 2.x, estas opciones de configuración tienen nombres de paquete diferentes. Para migrar a 2.x:

1. Cambie las instrucciones de importación de `com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibCon` y `com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel` a `software.amazon.kinesis.metrics.MetricsLevel` y `software.amazon.kinesis.metrics.MetricsUtil` respectivamente.

   ```
   // import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel
   import software.amazon.kinesis.metrics.MetricsLevel

   // import
    com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
   import software.amazon.kinesis.metrics.MetricsUtil
   ```

2. Sustituya la línea `metricsEnabledDimensionsKinesisClientLibConfiguration.DEFAULT_METRICS_ENABL` por `metricsEnabledDimensionsSet(MetricsUtil.OPERATION_DIMENSION_NAME, MetricsUtil.SHARD_ID_DIMENSION_NAME)`

A continuación, se presenta una versión actualizada de `KinesisInputDStream` con dimensiones y niveles de métricas personalizados:

```
import software.amazon.kinesis.metrics.MetricsLevel
import software.amazon.kinesis.metrics.MetricsUtil

val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
```

```
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)
  .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
 MetricsUtil.SHARD_ID_DIMENSION_NAME))
  .build()
```

- Función del controlador de mensajes en `KinesisInputDStream`

  Al crear una instancia de `KinesisInputDStream`, también puede proporcionar una "función del controlador de mensajes" que tome un registro de Kinesis y devuelva un objeto genérico T, en caso de que desee utilizar otros datos incluidos en un registro, como la clave de partición.

  En KCL 1.x, la firma de la función del controlador de mensajes es: `Record => T`, donde Record es `com.amazonaws.services.kinesis.model.Record`. En KCL 2.x, la firma del controlador se cambia a:`KinesisClientRecord => T`, where is. KinesisClientRecord `software.amazon.kinesis.retrieval.KinesisClientRecord`

  A continuación, se muestra un ejemplo de cómo proporcionar un controlador de mensajes en KCL 1.x:

```
import com.amazonaws.services.kinesis.model.Record


def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

  Para migrar el controlador de mensajes:

  1. Cambie la instrucción de importación de `com.amazonaws.services.kinesis.model.Record` a `software.amazon.kinesis.retrieval.KinesisClientRecord`.

```
// import com.amazonaws.services.kinesis.model.Record
import software.amazon.kinesis.retrieval.KinesisClientRecord
```

2. Actualice la firma del método del controlador de mensajes.

```
//def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt
 + 5
```

A continuación, se presenta un ejemplo actualizado de cómo proporcionar el controlador de mensajes en KCL 2.x:

```
import software.amazon.kinesis.retrieval.KinesisClientRecord


def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

Para obtener información sobre cómo migrar de KCL 1.x a KCL 2.x, consulte Migración de consumidores de KCL 1.x a KCL 2.x.

Migración de proveedores de AWS credenciales del SDK 1.x al AWS 2.x

Los proveedores de credenciales se utilizan para obtener AWS credenciales con las que interactuar. AWS Se han realizado varios cambios en las interfaces y clases relacionadas con los proveedores de credenciales en el SDK 2.x. Puede encontrar más información aquí. El conector Spark Kinesis ha definido una interfaz (`org.apache.spark.streaming.kinesis.SparkAWSCredentials`) y clases de implementación que devuelven la versión 1.x de los proveedores de AWS credenciales. Estos proveedores de credenciales son necesarios para inicializar los clientes de Kinesis. Por

ejemplo, si utiliza este método `SparkAWSCredentials.provider` en las aplicaciones, necesitará actualizar los códigos para utilizar la versión 2.x de los proveedores de credenciales. AWS

A continuación se muestra un ejemplo del uso de los proveedores de credenciales en AWS el SDK 1.x:

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import com.amazonaws.auth.AWSCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
    .basicCredentials("accessKey", "secretKey")
    .build()

val credentialProvider = basicSparkCredentials.provider
assert(credentialProvider.isInstanceOf[AWSCredentialsProvider], "Type should be
 AWSCredentialsProvider")
```

Para migrar a SDK 2.x:

1. Cambie la instrucción de importación de `com.amazonaws.auth.AWSCredentialsProvider` a `software.amazon.awssdk.auth.credentials.AwsCredentialsProvider`

   ```
   //import com.amazonaws.auth.AWSCredentialsProvider
   import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider
   ```

2. Actualice los códigos restantes que utilizan esta clase.

   ```
   import org.apache.spark.streaming.kinesis.SparkAWSCredentials
   import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider

   val basicSparkCredentials = SparkAWSCredentials.builder
       .basicCredentials("accessKey", "secretKey")
       .build()

   val credentialProvider = basicSparkCredentials.provider
   assert (credentialProvider.isInstanceOf[AwsCredentialsProvider], "Type should be
    AwsCredentialsProvider")
   ```

Migración de clientes de AWS servicios del AWS SDK 1.x al 2.x

AWS Los clientes de servicio tienen diferentes nombres de paquete en la versión 2.x (es decir,`software.amazon.awssdk`), mientras que en el SDK 1.x los usa. `com.amazonaws` Para obtener más información sobre los cambios de los clientes, consulte aquí. Si utiliza estos clientes de servicio en los códigos, deberá realizar la migración correspondiente de los clientes.

A continuación, se muestra un ejemplo de cómo crear un cliente en SDK 1.x:

```
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
import com.amazonaws.services.dynamodbv2.document.DynamoDB

AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();
```

Para migrar a 2.x:

1. Cambie las instrucciones de importación para los clientes de servicio. Tomemos como ejemplo los clientes de DynamoDB. Tendría que cambiar `com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient` o `com.amazonaws.services.dynamodbv2.document.DynamoDB` a `software.amazon.awssdk.services.dynamodb.DynamoDbClient`.

   ```
   // import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
   // import com.amazonaws.services.dynamodbv2.document.DynamoDB
   import software.amazon.awssdk.services.dynamodb.DynamoDbClient
   ```

2. Actualización de los códigos que inicializan los clientes

   ```
   // AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
   // AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();

   DynamoDbClient ddbClient = DynamoDbClient.create();
   DynamoDbClient ddbClient = DynamoDbClient.builder().build();
   ```

   Para obtener más información sobre la migración del AWS SDK de la versión 1.x a la 2.x, consulte ¿Qué diferencias hay entre el AWS SDK para Java 1.x y 2.x?

Ejemplos de código para aplicaciones de transmisión

```
import java.net.URI
import software.amazon.awssdk.auth.credentials.DefaultCredentialsProvider
import software.amazon.awssdk.http.apache.ApacheHttpClient
import software.amazon.awssdk.services.kinesis.KinesisClient
import software.amazon.awssdk.services.kinesis.model.DescribeStreamRequest
import software.amazon.awssdk.regions.Region
import software.amazon.kinesis.metrics.{MetricsLevel, MetricsUtil}

import org.apache.spark.SparkConf
import org.apache.spark.storage.StorageLevel
import org.apache.spark.streaming.{Milliseconds, StreamingContext}
import org.apache.spark.streaming.dstream.DStream.toPairDStreamFunctions
import org.apache.spark.streaming.kinesis.KinesisInitialPositions.Latest
import org.apache.spark.streaming.kinesis.KinesisInputDStream


object KinesisWordCountASLSDKV2 {

  def main(args: Array[String]): Unit = {
    val appName = "demo-app"
    val streamName = "demo-kinesis-test"
    val endpointUrl = "https://kinesis.us-west-2.amazonaws.com"
    val regionName = "us-west-2"

    // Determine the number of shards from the stream using the low-level Kinesis
 Client
    // from the AWS Java SDK.
    val credentialsProvider = DefaultCredentialsProvider.create
    require(credentialsProvider.resolveCredentials() != null,
      "No AWS credentials found. Please specify credentials using one of the methods
 specified " +
        "in https://docs.aws.amazon.com/sdk-for-java/latest/developer-guide/
credentials.html")
    val kinesisClient = KinesisClient.builder()
      .credentialsProvider(credentialsProvider)
      .region(Region.US_WEST_2)
      .endpointOverride(URI.create(endpointUrl))
      .httpClientBuilder(ApacheHttpClient.builder())
      .build()
    val describeStreamRequest = DescribeStreamRequest.builder()
      .streamName(streamName)
      .build()
```

```
    val numShards = kinesisClient.describeStream(describeStreamRequest)
      .streamDescription
      .shards
      .size


    // In this example, we are going to create 1 Kinesis Receiver/input DStream for
each shard.
    // This is not a necessity; if there are less receivers/DStreams than the number of
shards,
    // then the shards will be automatically distributed among the receivers and each
receiver
    // will receive data from multiple shards.
    val numStreams = numShards

    // Spark Streaming batch interval
    val batchInterval = Milliseconds(2000)

    // Kinesis checkpoint interval is the interval at which the DynamoDB is updated
with information
    // on sequence number of records that have been received. Same as batchInterval for
this
    // example.
    val kinesisCheckpointInterval = batchInterval

    // Setup the SparkConfig and StreamingContext
    val sparkConfig = new SparkConf().setAppName("KinesisWordCountASLSDKV2")
    val ssc = new StreamingContext(sparkConfig, batchInterval)

    // Create the Kinesis DStreams
    val kinesisStreams = (0 until numStreams).map { i =>
      KinesisInputDStream.builder
        .streamingContext(ssc)
        .streamName(streamName)
        .endpointUrl(endpointUrl)
        .regionName(regionName)
        .initialPosition(new Latest())
        .checkpointAppName(appName)
        .checkpointInterval(kinesisCheckpointInterval)
        .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
        .metricsLevel(MetricsLevel.DETAILED)
        .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
MetricsUtil.SHARD_ID_DIMENSION_NAME))
        .build()
```

```
    }

    // Union all the streams
    val unionStreams = ssc.union(kinesisStreams)

    // Convert each line of Array[Byte] to String, and split into words
    val words = unionStreams.flatMap(byteArray => new String(byteArray).split(" "))

    // Map each word to a (word, 1) tuple so we can reduce by key to count the words
    val wordCounts = words.map(word => (word, 1)).reduceByKey(_ + _)

    // Print the first 10 wordCounts
    wordCounts.print()

    // Start the streaming context and await termination
    ssc.start()
    ssc.awaitTermination()
  }
}
```

Consideraciones a la hora de utilizar el conector de Spark Kinesis actualizado

- Si sus aplicaciones utilizan la `Kinesis-producer-library` con una versión del JDK inferior a la 11, es posible que encuentre excepciones como `java.lang.NoClassDefFoundError: javax/xml/bind/DatatypeConverter`. Esto sucede porque EMR 7.0 incluye JDK 17 de forma predeterminada y los módulos J2EE se han eliminado de las bibliotecas estándar desde Java 11 en adelante. Esto se puede solucionar si agrega la siguiente dependencia en el archivo pom. Reemplace la versión de biblioteca por la que crea conveniente.

```
<dependency>
      <groupId>javax.xml.bind</groupId>
      <artifactId>jaxb-api</artifactId>
      <version>${jaxb-api.version}</version>
    </dependency>
```

- El jar del conector de Spark Kinesis se encuentra en esta ruta después de crear un clúster de EMR: `/usr/lib/spark/connector/lib/`

S3 (s3-dist-cp) DistCp

Apache DistCp es una herramienta de código abierto que puede utilizar para copiar grandes cantidades de datos. S3 DistCp es similar a Amazon S3 DistCp, pero está optimizado para funcionar con él AWS, especialmente. El comando para S3 DistCp en Amazon EMR versión 4.0 y versiones posteriores es `s3-dist-cp` el que se agrega como un paso en un clúster o en la línea de comandos. Con S3DistCp, puede copiar de manera eficiente grandes cantidades de datos de Amazon S3 a HDFS, donde se pueden procesar en los siguientes pasos de su clúster de Amazon EMR. También puede usar S3 DistCp para copiar datos entre buckets de Amazon S3 o desde HDFS a Amazon S3. S3 DistCp es más escalable y eficiente para copiar en paralelo grandes cantidades de objetos entre depósitos y AWS cuentas.

Para ver comandos específicos que demuestran la flexibilidad de S3DistCP en situaciones reales, consulte Siete consejos para usar S3 DistCp en el blog Big Data. AWS

Por ejemplo DistCp, S3 DistCp suele MapReduce copiar de forma distribuida. Comparte las tareas de copia, tratamiento de errores, recuperación y elaboración de informes entre varios servidores. Para obtener más información sobre el proyecto de código DistCp abierto Apache, consulte la DistCpguía en la documentación de Apache Hadoop.

Si S3 no puede copiar algunos o todos los archivos especificados, DistCp se produce un error en el paso del clúster y devuelve un código de error distinto de cero. Si esto ocurre, S3 DistCp no limpia los archivos parcialmente copiados.

> **Important**
>
> S3 DistCp no admite nombres de bucket de Amazon S3 que contengan el carácter de subrayado.
> S3 no DistCp admite la concatenación de archivos de Parquet. Úselo en su lugar. PySpark Para obtener más información, consulte Concatenación de archivos Parquet en Amazon EMR.
> Para evitar errores de copia al utilizar S3DistCP para copiar un único archivo (en lugar de un directorio) de S3 a HDFS, utilice Amazon EMR versión 5.33.0 o posterior, o Amazon EMR versión 6.3.0 o posterior.

DistCp Opciones de S3

Aunque es similar a DistCp, S3 DistCp admite un conjunto diferente de opciones para cambiar la forma en que copia y comprime los datos.

Al llamar a S3DistCp, puede especificar las opciones que se describen en la siguiente tabla. Las opciones se añaden al paso utilizando la lista de argumentos. En la siguiente tabla se muestran ejemplos de los DistCp argumentos de S3.

Opción	Descripción	Obligatorio
`--src=LOCATION`	Ubicación de los datos que copiar. Puede ser una ubicación de HDFS o Amazon S3. Ejemplo: `--src=s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node` **Important** S3 DistCp no admite nombres de bucket de Amazon S3 que contengan el carácter de subrayado.	Sí
`--dest=LOCATION`	Destino de los datos. Puede ser una ubicación de HDFS o Amazon S3. Ejemplo: `--dest=hdfs:///output` **Important** S3 DistCp no admite nombres de bucket de Amazon S3 que contengan el carácter de subrayado.	Sí
`--srcPattern=PATTERN`	Una expresión regular que filtra la operación de copia para un subconjunto de los datos de `--src`.	No

Opción	Descripción	Obligatorio
	Si no se especifica `--srcPattern` ni `--groupBy`, todos los datos de `--src` se copian en `--dest`. Si el argumento de expresión regular contiene caracteres especiales, como un asterisco (*), la expresión regular o toda la cadena `--args` debe estar escrita entre comillas ('). Ejemplo: `--srcPattern=.*daemons.*-hadoop-.*`	

Opción	Descripción	Obligatorio
`--groupBy=PATTERN`	Expresión regular que hace que S3 DistCp concatene los archivos que coinciden con la expresión. Por ejemplo, podría utilizar esta opción para combinar todos los archivos de registro escritos en una hora en un solo archivo. El nombre de archivo concatenado es el valor asignado por la expresión regular para la agrupación. Los paréntesis indican cómo deberían agruparse los archivos, todos los elementos que coincidan con la instrucción parentética se combinan en un único archivo de salida. Si la expresión regular no incluye una sentencia entre paréntesis, el clúster falla en el DistCp paso S3 y devuelve un error. Si el argumento de expresión regular contiene caracteres especiales, como un asterisco (*), la expresión regular o toda la cadena `--args` debe estar escrita entre comillas ('). Cuando se especifica `--groupBy` , solo se copian los archivos que coinciden con el patrón especificado. No tiene que especificar `--groupBy` y `--srcPattern` al mismo tiempo. Ejemplo: `--groupBy=.*subnetid.*([0-9]+-[0-9]+-[0-9]+-[0-9]+).*`	No

Opción	Descripción	Obligatorio
`--targetSize=SIZE`	El tamaño, en mebibytes (MiB), de los archivos a crear en función de la opción `--groupBy` . Este valor debe ser un número entero. Cuando `--targetSize` se establece, S3 DistCp intenta igualar este tamaño; el tamaño real de los archivos copiados puede ser mayor o menor que este valor. Los trabajos se agregan en función del tamaño del archivo de datos, por lo que es posible que el tamaño del archivo de destino coincida con el tamaño del archivo de datos de origen. Si los archivos concatenados por `--groupBy` son mayores que el valor de `--targetSize` , se dividen en archivos parciales y se nombran de forma secuencial con un valor numérico junto al final. Por ejemplo, un archivo concatenado en `myfile.gz` se desglosaría en partes como: `myfile0.gz` , `myfile1.gz` , etc. Ejemplo: `--targetSize=2`	No
`--appendToLastFile`	Especifica el comportamiento de S3 DistCp al copiar archivos de Amazon S3 a HDFS que ya están presentes. Añade los nuevos datos de archivo a archivos existentes. Si utiliza `--appendToLastFile` con `--groupBy` , se añaden nuevos datos a los archivos que coincidirán con los mismos grupos. Esta opción también respeta el comportamiento `--targetSize` cuando se utiliza con `--groupBy.`	No

Opción	Descripción	Obligatorio
`--outputCodec=CODEC`	Especifica el códec de compresión que utilizar para los archivos copiados. Esto puede tomar los valores: `gzip`, `gz`, `lzo`, `snappy` o bien `none`. Puede utilizar esta opción, por ejemplo, para convertir archivos de entrada comprimidos con Gzip en archivos de salida con compresión LZO o para descomprimir los archivos como parte de la operación de copia. Si elige un códec de salida, el nombre de archivo se agregará con la extensión adecuada (por ejemplo, `gz` y `gzip`, la extensión es `.gz`). Si no especifica un valor para `--outputCodec` , los archivos se copian sin cambios en su compresión. Ejemplo: `--outputCodec=lzo`	No
`--s3ServerSideEncryption`	Garantiza que los datos de destino se transfieran mediante SSL y se cifren automáticamente en Amazon S3 mediante una clave AWS del lado del servicio. Al recuperar datos mediante S3DistCp, los objetos se descifran automáticamente. Si intenta copiar un objeto sin cifrar en un bucket de Amazon S3 con cifrado requerido, la operación produce un error. Para obtener más información, consulte el tema relacionado con el uso del cifrado de datos. Ejemplo: `--s3ServerSideEncryption`	No

Opción	Descripción	Obligatorio
`--deleteOnSuccess`	Si la operación de copia se realiza correctamente, esta opción hace DistCp que S3 elimine los archivos copiados de la ubicación de origen. Esto resulta útil si está copiando archivos de salida, como archivos de registro, desde una ubicación a otra como una tarea programada y no desea copiar los mismos archivos dos veces. Ejemplo: `--deleteOnSuccess`	No
`--disableMultipartUpload`	Deshabilita el uso de la carga multiparte. Ejemplo: `--disableMultipartUpload`	No
`--multipartUploadChunkSize=SIZE`	El tamaño, en MiB, de cada parte de una carga multiparte de Amazon S3. S3 DistCp utiliza la carga multiparte cuando copia datos de un tamaño superior al`multipartUploadChunkSize` . Para mejorar el rendimiento de los trabajos, puede aumentar el tamaño de cada parte. El tamaño predeterminado es 128 MB. Ejemplo: `--multipartUploadChunkSize=1000`	No
`--numberFiles`	Añade archivos de salida con los números secuenciales. El recuento comienza en 0 a menos que se especifique un valor diferente mediante `--startingIndex` . Ejemplo: `--numberFiles`	No
`--startingIndex=INDEX`	Utilizado con `--numberFiles` para especificar el primer número de la secuencia. Ejemplo: `--startingIndex=1`	No

Opción	Descripción	Obligatorio
`--outputManifest=FILENAME`	Crea un archivo de texto, comprimido con Gzip, que contiene una lista de todos los archivos copiados por S3. DistCp Ejemplo: `--outputManifest=manifest-1.gz`	No
`--previousManifest=PATH`	Lee un archivo de manifiesto que se creó durante una llamada anterior a S3 DistCp con la `--outputManifest` marca. Cuando el `--previousManifest` indicador está activado, S3 DistCp excluye de la operación de copia los archivos que figuran en el manifiesto. Si se especifica `--outputManifest` junto con los archivos `--previousManifest`, incluidos en el manifiesto anterior también aparecen en el nuevo archivo de manifiesto, aunque los archivos no se copiarán. Ejemplo: `--previousManifest=/usr/bin/manifest-1.gz`	No
`--requirePreviousManifest`	Requiere un manifiesto previo creado durante una llamada anterior a S3DistCp. Si se define como false, no se genera ningún error cuando no se especifica un manifiesto anterior. El valor predeterminado es true.	No
`--copyFromManifest`	Invierte el comportamiento de `--previousManifest` hacer que S3 DistCp utilice el archivo de manifiesto especificado como una lista de archivos para copiar, en lugar de una lista de archivos para excluirlos de la copia. Ejemplo: `--copyFromManifest --previousManifest=/usr/bin/manifest-1.gz`	No

Opción	Descripción	Obligatorio
`--s3Endpoint=ENDPOINT`	Especifica el punto de conexión de Amazon S3 que utilizar a la hora de cargar un archivo. Esta opción establece el punto de enlace para el origen y el destino. Si no se define, el punto de enlace predeterminado es `s3.amazonaws.com` . Para ver una lista de los puntos de conexión de Amazon S3, consulte Regiones y puntos de conexión. Ejemplo: `--s3Endpoint=s3.eu-west-1.amazonaws.com`	No
`--storageClass=CLASS`	La clase de almacenamiento que utilizar cuando el destino es Amazon S3. Los valores válidos son STANDARD y REDUCED_REDUNDANCY. Si no se especifica esta opción, S3 DistCp intenta conservar la clase de almacenamiento. Ejemplo: `--storageClass=STANDARD`	No

Opción	Descripción	Obligatorio
`--srcPrefixesFile=PATH`	un archivo de texto en Amazon S3 (s3://), HDFS (hdfs:///) o sistema de archivos local (file:/) que contiene una lista de prefijos `src`, un prefijo por línea. Si `srcPrefixesFile` se proporciona, S3 no DistCp mostrará la ruta src. En lugar de ello, genera una lista de origen como resultado combinado de enumerar todos los prefijos especificados en este archivo. La ruta relativa en comparación con la ruta src, en lugar de estos prefijos, se utilizará para generar las rutas de destino. Si también se especific a `srcPattern` , se aplicará a los resultados de lista combinada de los prefijos de origen para filtrar aún más la entrada. Si se utiliza `copyFromM anifest` , lo objetos en el manifiesto se copiarán y no se tendrá en cuenta `srcPrefixesFile` . Ejemplo: `--srcPrefixesFile=PATH`	No

Además de las opciones anteriores, S3 DistCp implementa la interfaz Tool, lo que significa que admite las opciones genéricas.

Añadir S3 DistCp como un paso de un clúster

Para llamar a S3, DistCp agréguelo como un paso en su clúster. Los pasos se pueden añadir a un clúster en el momento de su lanzamiento o a un clúster en ejecución mediante la consola, la CLI o la API. En los siguientes ejemplos, se muestra cómo añadir un DistCp paso de S3 a un clúster en ejecución. Para obtener más información sobre cómo agregar pasos a un clúster, consulte Enviar trabajo a un clúster en la Guía de administración de Amazon EMR.

Para añadir un DistCp paso S3 a un clúster en ejecución mediante el AWS CLI

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte la Referencia de AWS CLI comandos.

- Para añadir un paso a un clúster que llame a S3DistCp, pase los parámetros que especifican cómo S3 DistCp debe realizar la operación de copia como argumentos.

  El siguiente ejemplo copia registros de daemon desde Amazon S3 a `hdfs:///output`. En el comando siguiente:

  - `--cluster-id` especifica el clúster
  - `Jar`es la ubicación del archivo DistCp JAR de S3. Para ver un ejemplo de cómo ejecutar un comando en un clúster mediante command-runner.jar, consulte Envío de un paso JAR personalizado para ejecutar un script o un comando.
  - `Args`es una lista separada por comas de los pares nombre-valor de las opciones que se van a pasar a S3. DistCp Para ver una lista completa de las opciones disponibles, consulte DistCp Opciones de S3.

  Para añadir un paso de DistCp copia de S3 a un clúster en ejecución, coloque lo siguiente en un archivo JSON guardado en Amazon S3 o en su sistema de archivos local, como en este *`myStep.json`* ejemplo. *`j-3GYXXXXXX9IOK`*Sustitúyalo por tu ID de clúster y *`amzn-s3-demo-bucket`* sustitúyelo por el nombre de tu bucket de Amazon S3.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-
demo-bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*[a-
zA-Z,]+"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
]
```

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK --steps file://./myStep.json
```

Example Copie los archivos de registro desde Amazon S3 a HDFS

Este ejemplo también muestra cómo copiar archivos de registro almacenados en un bucket de Amazon S3 en HDFS agregando un paso a un clúster en ejecución. En este ejemplo, la opción `--srcPattern` se utiliza para limitar los datos copiados en los registros del daemon.

Para copiar los archivos de registro desde Amazon S3 en HDFS mediante la opción `--srcPattern`, coloque lo siguiente en un archivo JSON guardado en Amazon S3 o en el sistema de archivos local, como *`myStep.json`* en este ejemplo. *`j-3GYXXXXXX9IOK`*Sustitúyalo por tu ID de clúster y *`amzn-s3-demo-bucket`* sustitúyelo por el nombre de tu bucket de Amazon S3.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-demo-
bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*daemons.*-
hadoop-.*"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
]
```

Limpiando después de DistCp trabajos fallidos en S3

Si S3 DistCp no puede copiar algunos o todos los archivos especificados, el comando o el paso del clúster fallará y devolverá un código de error distinto de cero. Si esto ocurre, S3 DistCp no limpia los archivos parcialmente copiados. Deben eliminarse manualmente.

Los archivos copiados parcialmente se guardan en el `tmp` directorio HDFS, en subdirectorios, con el identificador único del trabajo de S3DistCp . Encontrará este ID en la salida estándar del trabajo.

Por ejemplo, para un DistCp trabajo de S3 con el ID`4b1c37bb-91af-4391-aaf8-46a6067085a6`, puede conectarse al nodo principal del clúster y ejecutar el siguiente comando para ver los archivos de salida asociados al trabajo.

```
hdfs dfs -ls /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output
```

El comando devuelve una lista de archivos similar a la siguiente:

```
Found 8 items
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/_SUCCESS
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00000
```

```
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00001
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00002
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00003
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00004
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00005
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00006
```

A continuación, puede ejecutar el siguiente comando para eliminar el directorio y todo el contenido.

```
hdfs dfs rm -rf /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6
```

Ejecución de comandos y scripts en un clúster de Amazon EMR

En este tema se explica cómo ejecutar un comando o un script como paso en el clúster. Ejecutar un comando o un script como paso es una de las muchas maneras de enviar el trabajo a un clúster y es útil en las siguientes situaciones:

- Cuando no tiene acceso SSH a su clúster de Amazon EMR
- Cuando desea ejecutar un comando bash o de intérprete de comandos para solucionar los problemas de su clúster

Puede ejecutar un script tanto al crear un clúster como si el clúster está en estado `WAITING`. Para ejecutar un script antes de que comience el procesamiento de un paso, utilice una acción de arranque en su lugar. Para obtener más información acerca de las acciones de arranque, consulte Creación de acciones de arranque para instalar software adicional en la Guía de administración de Amazon EMR.

Amazon EMR proporciona las siguientes herramientas para ayudarle a ejecutar scripts, comandos y otros programas en el clúster. Puede invocar ambas herramientas mediante la consola de administración de Amazon EMR o la AWS CLI.

`command-runner.jar`

Se encuentra en la AMI de Amazon EMR de su clúster. Puede utilizar `command-runner.jar` para ejecutar comandos en su clúster. Especifica `command-runner.jar` sin utilizar su ruta completa.

`script-runner.jar`

Se aloja en Amazon S3 en `s3://<region>.elasticmapreduce/libs/script-runner/script-runner.jar`, donde *`<region>`* se encuentra la región en la que reside su clúster de Amazon EMR. Puede usar `script-runner.jar` para ejecutar scripts guardados localmente o en Amazon S3 en su clúster. Debe especificar el URI completo de `script-runner.jar` al enviar un paso.

Envío de un paso JAR personalizado para ejecutar un script o un comando

Los siguientes AWS CLI ejemplos ilustran algunos casos de uso comunes de `command-runner.jar` y `script-runner.jar` en Amazon EMR.

Example : ejecutar un comando en un clúster mediante **`command-runner.jar`**

Cuando usa `command-runner.jar`, especifica comandos, opciones y valores en la lista de argumentos del paso.

El siguiente AWS CLI ejemplo envía un paso a un clúster en ejecución que lo invoca. `command-runner.jar` El comando especificado en la `Args` lista descarga un script llamado *`my-script.sh`* desde Amazon S3 al directorio principal del usuario de hadoop. A continuación, el comando modifica los permisos del script y se ejecuta. *`my-script.sh`*

Al usar el AWS CLI, los elementos de la `Args` lista deben estar separados por comas, sin espacios en blanco entre los elementos de la lista. Por ejemplo, `Args=[example-command,example-option,"example option value"]` en lugar de `Args=[example-command, example-option, "example option value"]`.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Download a script from S3, change its permissions, and
 run it",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[bash,-c,"aws s3 cp s3://
amzn-s3-demo-bucket/my-script.sh /home/hadoop; chmod u+x /home/hadoop/my-script.sh;
 cd /home/hadoop; ./my-script.sh"]
```

Example : Ejecutar un script en un clúster mediante **`script-runner.jar`**

Cuando usa `script-runner.jar`, especifica el script que desea ejecutar en la lista de argumentos de su paso.

En el siguiente AWS CLI ejemplo, se envía un paso a un clúster en ejecución que lo invoca. `script-runner.jar` En este caso, el script llamado *`my-script.sh`* se almacena en Amazon S3. También puede especificar los scripts locales que se almacenan en el nodo maestro del clúster.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
```

```
--steps Type=CUSTOM_JAR,Name="Run a script from S3 with script-
runner.jar",ActionOnFailure=CONTINUE,Jar=s3://us-west-2.elasticmapreduce/libs/script-
runner/script-runner.jar,Args=[s3://amzn-s3-demo-bucket/my-script.sh]
```

Otras formas de usar `command-runner.jar`

También se puede utilizar `command-runner.jar` para enviar el trabajo a un clúster con herramientas como `spark-submit` o `hadoop-streaming`. Al lanzar una aplicación con `command-runner.jar`, se especifica `CUSTOM_JAR` como tipo de paso en lugar de utilizar un valor como `SPARK`, `STREAMING` o `PIG`. La disponibilidad de las herramientas varía en función de las aplicaciones que haya instalado en el clúster.

El siguiente comando de ejemplo utiliza `command-runner.jar` para enviar un paso mediante `spark-submit`. La `Args` lista se especifica `spark-submit` como comando, seguido del URI de Amazon S3 de la aplicación Spark *`my-app.py`* con argumentos y valores.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Run spark-submit using command-
runner.jar",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[spark-submit,S3://
amzn-s3-demo-bucket/my-app.py,ArgName1,ArgValue1,ArgName2,ArgValue2]
```

En la siguiente tabla se identifican las herramientas adicionales que puede utilizar para ejecutar `command-runner.jar`.

Nombre de la herramienta	Descripción
`hadoop-streaming`	Envía un programa de streaming de Hadoop. En la consola y en algunas otras SDKs, se trata de un paso de streaming.
`hive-script`	Ejecuta un script de Hive. En la consola y SDKs, este es un paso de Hive.
`pig-script`	Ejecuta un script de Pig. En la consola y SDKs, este es un Pig Step.
`spark-submit`	Ejecuta una aplicación de Spark. En la consola, este es un paso de Spark.

Nombre de la herramienta	Descripción
`hadoop-lzo`	Ejecuta el indizador Hadoop LZO en un directorio.
`s3-dist-cp`	Copia distribuida de grandes cantidades de datos desde Amazon S3 en HDFS. Para obtener más información, consulte S3 (s3-dist-cp) DistCp .

AWS Glosario

Para obtener la AWS terminología más reciente, consulte el AWS glosario de la Glosario de AWS Referencia.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la version original de inglés, prevalecerá la version en inglés.