Cloudera Runtime 7.2.16

Release Notes

Date published: 2023-01-11
Date modified: 2024-02-05

https://docs.cloudera.com/

Legal Notice

Contents

Overview

You can review the Release Notes of Cloudera Runtime 7.2.16 for release-specific information related to new features and improvements, bug fixes, deprecated features and components, known issues, and changed features that can affect product behavior.

Cloudera Runtime Component Versions

You must be familiar with the versions of all the components in the Cloudera Runtime 7.2.16 distribution to ensure compatibility of these components with other applications. You must also be aware of the available Technical Preview components and use them only in a testing environment.

Apache Components

Component	Version
Apache Arrow	0.11.1.7.2.16.0-287
Apache Atlas	2.2.0.7.2.16.0-287
Apache Calcite	1.21.0.7.2.16.0-287
Apache Avro	1.8.2.7.2.16.0-287
Apache Flink	1.15.1.1.9.0.0
Apache Hadoop (Includes YARN and HDFS)	3.1.1.7.2.16.0-287
Apache HBase	2.4.6.7.2.16.0-287
Apache Hive	3.1.3000.7.2.16.0-287
Apache Iceberg	0.14.1
Apache Impala	4.0.0.7.2.16.0-287
Apache Kafka	3.1.2.7.2.16.0-287
Apache Knox	1.3.0.7.2.16.0-287
Apache Kudu	1.15.0.7.2.16.0-287
Apache Livy	0.6.0.7.2.16.0-287
Apache MapReduce	3.1.1.7.2.16.0-287
Apache NiFi	1.18.0.2.2.6.0
Apache NiFi Registry	1.18.0.2.2.6.0
Apache Oozie	5.1.0.7.2.16.0-287
Apache ORC	1.5.1.7.2.16.0-287
Apache Parquet	1.10.99.7.2.16.0-287
Apache Phoenix	5.1.1.7.2.16.0-287
Apache Ranger	2.3.0.7.2.16.0-287
Apache Solr	8.4.1.7.2.16.0-287
Apache Spark	2.4.8.7.2.16.0-287
Apache Spark 3	3.3.0.7.2.16.0-287
Apache Sqoop	1.4.7.7.2.16.0-287
Apache Tez	0.9.1.7.2.16.0-287

Component	Version
Apache Zeppelin	0.8.2.7.2.16.0-287
Apache ZooKeeper	3.5.5.7.2.16.0-287

Other Components

Component	Version
Cruise Control	2.5.85.7.2.16.0-551
Data Analytics Studio	1.4.2.7.2.16.0-287
GCS Connector	2.1.2.7.2.16.0-287
HBase Indexer	1.5.0.7.2.16.0-287
Hive Solr Connector	4.0.0.7.2.16.0-287
Hue	4.5.0.7.2.16.0-287
Search	1.0.0.7.2.16.0-287
Schema Registry	0.10.0.7.2.16.0-287
Spark Solr Connector	3.9.0.7.2.16.0-287
Streams Messaging Manager	2.3.0.7.2.16.0-287
Streams Replication Manager	1.1.0.7.2.16.0-287

Connectors and Encryption Components

Component	Version
HBase connectors	1.0.0.7.2.16.0-287
Hive Meta Store (HMS)	1.0.0.7.2.16.0-287
Hive on Tez	1.0.0.7.2.16.0-287
Hive Warehouse Connector	1.0.0.7.2.16.0-287
Spark Atlas Connector	0.1.0.7.2.16.0-287
Spark Schema Registry	1.1.0.7.2.16.0-287

Using the Cloudera Runtime Maven repository 7.2.16

Information about using Maven to build applications with Cloudera Runtime components.

If you want to build applications or tools for use with Cloudera Runtime components and you are using Maven or Ivy for dependency management, you can pull the Cloudera Runtime artifacts from the Cloudera Maven repository. The repository is available at https://repository.cloudera.com/artifactory/cloudera-repos/.

Important: When you build an application JAR, do not include CDH JARs, because they are already provided. If you do, upgrading CDH can break your application. To avoid this situation, set the Maven dependency scope to provided. If you have already built applications which include the CDH JARs, update the dependency to set scope to provided and recompile.

The following is a sample POM (pom.xml) file:

```
<project xmlns="http://maven.apache.org/POM/4.0.0" xmlns:xsi="http://www.w3.
org/2001/XMLSchema-instance" xsi:schemaLocation="http://maven.apache.org/POM
/4.0.0 http://maven.apache.org/maven-v4_0_0.xsd">
  <repositories>
    <repository>
```

```
      <id>cloudera</id>
      <url>https://repository.cloudera.com/artifactory/cloudera-repos/</url>
    </repository>
  </repositories>
</project>
```

Maven Artifacts for Cloudera Runtime 7.2.16

The following table lists the project name, groupId, artifactId, and version required to access each RUNTIME artifact.

Project	groupId	artifactId	version
Apache Atlas	org.apache.atlas	atlas-authorization	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-aws-s3-bridge	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-azure-adls-bridge	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-classification-updater	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-client-common	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-client-v1	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-client-v2	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-common	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-distro	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-docs	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-graphdb-api	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-graphdb-common	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-graphdb-janus	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-hdfs-bridge	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-index-repair-tool	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-intg	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-janusgraph-hbase2	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-notification	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-plugin-classloader	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-repository	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-server-api	2.1.0.7.2.16.0-287
	org.apache.atlas	atlas-testtools	2.1.0.7.2.16.0-287
	org.apache.atlas	hbase-bridge	2.1.0.7.2.16.0-287
	org.apache.atlas	hbase-bridge-shim	2.1.0.7.2.16.0-287
	org.apache.atlas	hbase-testing-util	2.1.0.7.2.16.0-287
	org.apache.atlas	hdfs-model	2.1.0.7.2.16.0-287
	org.apache.atlas	hive-bridge	2.1.0.7.2.16.0-287
	org.apache.atlas	hive-bridge-shim	2.1.0.7.2.16.0-287
	org.apache.atlas	impala-bridge	2.1.0.7.2.16.0-287
	org.apache.atlas	impala-bridge-shim	2.1.0.7.2.16.0-287
	org.apache.atlas	impala-hook-api	2.1.0.7.2.16.0-287
	org.apache.atlas	kafka-bridge	2.1.0.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.atlas	kafka-bridge-shim	2.1.0.7.2.16.0-287
	org.apache.atlas	navigator-to-atlas	2.1.0.7.2.16.0-287
	org.apache.atlas	sample-app	2.1.0.7.2.16.0-287
	org.apache.atlas	sqoop-bridge	2.1.0.7.2.16.0-287
	org.apache.atlas	sqoop-bridge-shim	2.1.0.7.2.16.0-287
Apache Avro	org.apache.avro	avro	1.8.2.7.2.16.0-287
	org.apache.avro	avro-compiler	1.8.2.7.2.16.0-287
	org.apache.avro	avro-ipc	1.8.2.7.2.16.0-287
	org.apache.avro	avro-mapred	1.8.2.7.2.16.0-287
	org.apache.avro	avro-maven-plugin	1.8.2.7.2.16.0-287
	org.apache.avro	avro-protobuf	1.8.2.7.2.16.0-287
	org.apache.avro	avro-service-archetype	1.8.2.7.2.16.0-287
	org.apache.avro	avro-thrift	1.8.2.7.2.16.0-287
	org.apache.avro	avro-tools	1.8.2.7.2.16.0-287
	org.apache.avro	trevni-avro	1.8.2.7.2.16.0-287
	org.apache.avro	trevni-core	1.8.2.7.2.16.0-287
Apache Calcite	org.apache.calcite	calcite-babel	1.21.0.7.2.16.0-287
	org.apache.calcite	calcite-core	1.21.0.7.2.16.0-287
	org.apache.calcite	calcite-druid	1.21.0.7.2.16.0-287
	org.apache.calcite	calcite-kafka	1.21.0.7.2.16.0-287
	org.apache.calcite	calcite-linq4j	1.21.0.7.2.16.0-287
	org.apache.calcite	calcite-server	1.21.0.7.2.16.0-287
	org.apache.calcite	calcite-ubenchmark	1.21.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica	1.17.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica-core	1.17.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica-metrics	1.17.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica-metrics-dropwizardmetrics	1.17.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica-noop-driver	1.17.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica-server	1.17.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica-standalone-server	1.17.0.7.2.16.0-287
	org.apache.calcite.avatica	avatica-tck	1.17.0.7.2.16.0-287
GCS Connector	com.google.cloud.bigdataoss	bigquery-connector	2.1.2.7.2.16.0-287
	com.google.cloud.bigdataoss	gcs-connector	2.1.2.7.2.16.0-287
	com.google.cloud.bigdataoss	gcsio	2.1.2.7.2.16.0-287
	com.google.cloud.bigdataoss	util	2.1.2.7.2.16.0-287
	com.google.cloud.bigdataoss	util-hadoop	2.1.2.7.2.16.0-287
Apache Hadoop	org.apache.hadoop	hadoop-aliyun	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-annotations	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-archive-logs	3.1.1.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-archives	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-assemblies	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-auth	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-aws	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-azure	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-azure-datalake	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-benchmark	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-build-tools	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-client	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-client-api	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-client-integration-tests	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-client-minicluster	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-client-runtime	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-cloud-storage	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-common	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-datajoin	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-distcp	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-extras	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-fs2img	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-gridmix	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-hdfs	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-hdfs-client	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-hdfs-httpfs	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-hdfs-native-client	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-hdfs-nfs	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-hdfs-rbf	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-kafka	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-kms	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-app	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-common	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-core	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-hs	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-hs-plugins	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-jobclient	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-nativetask	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-shuffle	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-client-uploader	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-mapreduce-examples	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-maven-plugins	3.1.1.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-minicluster	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-minikdc	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-nfs	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-openstack	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-resourceestimator	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-rumen	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-sls	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-streaming	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-tools-dist	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-api	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-applications-distributedshell	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-applications-unmanaged-am-launcher	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-client	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-common	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-registry	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-applicationhistoryservice	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-common	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-nodemanager	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-resourcemanager	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-router	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-sharedcachemanager	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-tests	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-timeline-pluginstorage	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-timelineservice	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-client	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-common	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-server-2	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-tests	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-server-web-proxy	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-services-api	3.1.1.7.2.16.0-287
	org.apache.hadoop	hadoop-yarn-services-core	3.1.1.7.2.16.0-287
Apache HBase	org.apache.hbase	hbase-annotations	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-asyncfs	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-checkstyle	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-client	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-client-project	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-common	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-endpoint	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-examples	2.4.6.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.hbase	hbase-external-blockcache	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-hadoop-compat	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-hadoop2-compat	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-hbtop	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-http	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-it	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-logging	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-mapreduce	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-metrics	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-metrics-api	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-procedure	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-protocol	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-protocol-shaded	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-replication	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-resource-bundle	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-rest	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-rsgroup	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-server	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-shaded-client	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-shaded-client-byo-hadoop	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-shaded-client-project	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-shaded-mapreduce	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-shaded-testing-util	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-shaded-testing-util-tester	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-shell	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-testing-util	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-thrift	2.4.6.7.2.16.0-287
	org.apache.hbase	hbase-zookeeper	2.4.6.7.2.16.0-287
	org.apache.hbase.connectors.kafka	hbase-kafka-model	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.kafka	hbase-kafka-proxy	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark-it	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark-protocol	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark-protocol-shaded	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark3	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark3-it	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark3-protocol	1.0.0.7.2.16.0-287
	org.apache.hbase.connectors.spark	hbase-spark3-protocol-shaded	1.0.0.7.2.16.0-287
	org.apache.hbase.filesystem	hbase-hadoop-testutils	1.0.0.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.hbase.filesystem	hbase-filesystem-impl	1.0.0.7.2.16.0-287
	org.apache.hbase.filesystem	hbase-oss	1.0.0.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-noop-htrace	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-shaded-gson	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-shaded-jackson-jaxrs-json-provider	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-shaded-jersey	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-shaded-jetty	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-shaded-miscellaneous	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-shaded-netty	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-shaded-protobuf	4.1.1.7.2.16.0-287
	org.apache.hbase.thirdparty	hbase-unsafe	4.1.1.7.2.16.0-287
Apache Hive	org.apache.hive	catalogd-unit	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-beeline	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-blobstore	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-classification	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-cli	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-common	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-contrib	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-exec	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-hbase-handler	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-hcatalog-it-unit	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-hplsql	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-iceberg-catalog	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-iceberg-handler	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-iceberg-shading	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-impala	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-custom-serde	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-iceberg	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-impala	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-minikdc	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-qfile	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-qfile-kudu	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-test-serde	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-unit	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-unit-hadoop2	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-it-util	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-jdbc	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-jdbc-handler	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-jmh	3.1.3000.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.hive	hive-kudu-handler	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-llap-client	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-llap-common	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-llap-ext-client	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-llap-server	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-llap-tez	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-metastore	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-parser	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-pre-upgrade	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-serde	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-service	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-service-rpc	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-shims	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-standalone-metastore	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-storage-api	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-streaming	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-testutils	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-udf	3.1.3000.7.2.16.0-287
	org.apache.hive	hive-vector-code-gen	3.1.3000.7.2.16.0-287
	org.apache.hive	kafka-handler	3.1.3000.7.2.16.0-287
	org.apache.hive.hcatalog	hive-hcatalog-core	3.1.3000.7.2.16.0-287
	org.apache.hive.hcatalog	hive-hcatalog-pig-adapter	3.1.3000.7.2.16.0-287
	org.apache.hive.hcatalog	hive-hcatalog-server-extensions	3.1.3000.7.2.16.0-287
	org.apache.hive.hcatalog	hive-hcatalog-streaming	3.1.3000.7.2.16.0-287
	org.apache.hive.hcatalog	hive-webhcat	3.1.3000.7.2.16.0-287
	org.apache.hive.hcatalog	hive-webhcat-java-client	3.1.3000.7.2.16.0-287
	org.apache.hive.hive-it-custom-udfs	udf-classloader-udf1	3.1.3000.7.2.16.0-287
	org.apache.hive.hive-it-custom-udfs	udf-classloader-udf2	3.1.3000.7.2.16.0-287
	org.apache.hive.hive-it-custom-udfs	udf-classloader-util	3.1.3000.7.2.16.0-287
	org.apache.hive.hive-it-custom-udfs	udf-vectorized-badexample	3.1.3000.7.2.16.0-287
	org.apache.hive.shims	hive-shims-0.23	3.1.3000.7.2.16.0-287
	org.apache.hive.shims	hive-shims-common	3.1.3000.7.2.16.0-287
	org.apache.hive.shims	hive-shims-scheduler	3.1.3000.7.2.16.0-287
Apache Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector-spark3_2.12	1.0.0.7.2.16.0-287
	com.hortonworks.hive	hive-warehouse-connector_2.11	1.0.0.7.2.16.0-287
Apache Kafka	org.apache.kafka	connect	3.1.2.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.kafka	connect-api	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-basic-auth-extension	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-cloudera-authorization-extension	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-cloudera-common	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-cloudera-secret-storage	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-cloudera-security-policies	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-file	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-json	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-mirror	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-mirror-client	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-runtime	3.1.2.7.2.16.0-287
	org.apache.kafka	connect-transforms	3.1.2.7.2.16.0-287
	org.apache.kafka	generator	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-clients	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-cloudera-metrics-reporter_2.12	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-cloudera-metrics-reporter_2.13	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-cloudera-plugins	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-examples	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-log4j-appender	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-metadata	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-raft	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-server-common	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-shell	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-storage	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-storage-api	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-examples	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-scala_2.12	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-scala_2.13	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-test-utils	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-0100	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-0101	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-0102	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-0110	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-10	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-11	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-20	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-21	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-22	3.1.2.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.kafka	kafka-streams-upgrade-system-tests-23	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-24	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-25	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-26	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-27	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-28	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-streams-upgrade-system-tests-30	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka-tools	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka_2.12	3.1.2.7.2.16.0-287
	org.apache.kafka	kafka_2.13	3.1.2.7.2.16.0-287
	org.apache.kafka	trogdor	3.1.2.7.2.16.0-287
Apache Knox	org.apache.knox	gateway-adapter	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-admin-ui	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-applications	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-cloud-bindings	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-demo-ldap	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-demo-ldap-launcher	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-discovery-ambari	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-discovery-cm	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-docker	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-i18n	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-i18n-logging-log4j	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-i18n-logging-sl4j	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-performance-test	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-ha	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-identity-assertion-common	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-identity-assertion-concat	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-identity-assertion-hadoop-groups	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-identity-assertion-no-doas	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-identity-assertion-pseudo	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-identity-assertion-regex	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-identity-assertion-switchcase	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-jersey	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-rewrite	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-rewrite-common	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-rewrite-func-hostmap-static	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-rewrite-func-inbound-query-param	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-rewrite-func-service-registry	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-rewrite-step-encrypt-uri	1.3.0.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.knox	gateway-provider-rewrite-step-secure-query	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-authc-anon	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-authz-acls	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-authz-composite	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-clientcert	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-hadoopauth	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-jwt	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-pac4j	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-preauth	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-shiro	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-provider-security-webappsec	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-release	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-server	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-server-launcher	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-server-xforwarded-filter	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-admin	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-as	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-auth	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-definitions	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-hashicorp-vault	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-hbase	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-health	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-hive	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-idbroker	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-impala	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-jkg	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-knoxsso	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-knoxssout	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-knoxtoken	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-livy	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-metadata	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-nifi	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-nifi-registry	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-remoteconfig	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-rm	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-session	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-storm	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-test	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-tgs	1.3.0.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.knox	gateway-service-vault	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-service-webhdfs	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-shell	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-shell-launcher	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-shell-release	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-shell-samples	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-spi	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-spi-common	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-test	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-test-idbroker	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-test-release-utils	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-test-utils	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-topology-hadoop-xml	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-topology-simple	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-util-common	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-util-configinjector	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-util-launcher	1.3.0.7.2.16.0-287
	org.apache.knox	gateway-util-urltemplate	1.3.0.7.2.16.0-287
	org.apache.knox	hadoop-examples	1.3.0.7.2.16.0-287
	org.apache.knox	knox-cli-launcher	1.3.0.7.2.16.0-287
	org.apache.knox	knox-homepage-ui	1.3.0.7.2.16.0-287
	org.apache.knox	knox-token-management-ui	1.3.0.7.2.16.0-287
	org.apache.knox	knox-webshell-ui	1.3.0.7.2.16.0-287
	org.apache.knox	webhdfs-kerb-test	1.3.0.7.2.16.0-287
	org.apache.knox	webhdfs-test	1.3.0.7.2.16.0-287
Apache Kudu	org.apache.kudu	kudu-backup-tools	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-backup2_2.11	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-backup3_2.12	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-client	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-hive	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-spark2-tools_2.11	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-spark2_2.11	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-spark3-tools_2.12	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-spark3_2.12	1.15.0.7.2.16.0-287
	org.apache.kudu	kudu-test-utils	1.15.0.7.2.16.0-287
Apache Livy	org.apache.livy	livy-api	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-client-common	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-client-http	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-core_2.11	0.6.0.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.livy	livy-core_2.12	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-examples	0.6.0.7.2.16.0-287
	org.apache.livy	livy-integration-test	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-repl_2.11	0.6.0.7.2.16.0-287
	org.apache.livy	livy-repl_2.12	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-rsc	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-scala-api_2.11	0.6.0.7.2.16.0-287
	org.apache.livy	livy-scala-api_2.12	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-server	0.6.0.7.2.16.0-287
	org.apache.livy	livy-test-lib	0.6.0.7.2.16.0-287
	org.apache.livy	livy-thriftserver	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-thriftserver-session	0.6.3000.7.2.16.0-287
	org.apache.livy	livy-thriftserver-session	0.6.0.7.2.16.0-287
Apache Lucene	org.apache.lucene	lucene-analyzers-common	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-icu	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-kuromoji	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-morfologik	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-nori	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-opennlp	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-phonetic	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-smartcn	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-analyzers-stempel	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-backward-codecs	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-benchmark	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-classification	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-codecs	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-core	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-demo	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-expressions	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-facet	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-grouping	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-highlighter	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-join	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-memory	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-misc	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-monitor	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-queries	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-queryparser	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-replicator	8.4.1.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.lucene	lucene-sandbox	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-spatial	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-spatial-extras	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-spatial3d	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-suggest	8.4.1.7.2.16.0-287
	org.apache.lucene	lucene-test-framework	8.4.1.7.2.16.0-287
Apache Oozie	org.apache.oozie	oozie-client	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-core	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-distro	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-examples	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-fluent-job-api	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-fluent-job-client	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-server	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-distcp	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-git	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-hcatalog	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-hive	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-hive2	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-oozie	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-spark	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-sqoop	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-sharelib-streaming	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-tools	5.1.0.7.2.16.0-287
	org.apache.oozie	oozie-zookeeper-security-tests	5.1.0.7.2.16.0-287
	org.apache.oozie.test	oozie-mini	5.1.0.7.2.16.0-287
Apache ORC	org.apache.orc	orc-core	1.5.1.7.2.16.0-287
	org.apache.orc	orc-examples	1.5.1.7.2.16.0-287
	org.apache.orc	orc-mapreduce	1.5.1.7.2.16.0-287
	org.apache.orc	orc-shims	1.5.1.7.2.16.0-287
	org.apache.orc	orc-tools	1.5.1.7.2.16.0-287
Apache Parquet	org.apache.parquet	parquet-avro	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-cascading	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-cascading3	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-column	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-common	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-encoding	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-format-structures	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-generator	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-hadoop	1.10.99.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.parquet	parquet-hadoop-bundle	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-jackson	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-pig	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-pig-bundle	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-protobuf	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-scala_2.10	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-thrift	1.10.99.7.2.16.0-287
	org.apache.parquet	parquet-tools	1.10.99.7.2.16.0-287
Apache Phoenix	org.apache.phoenix	phoenix-client-embedded-hbase-2.4	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-client-hbase-2.4	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-connectors-phoenix5-compat	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix-core	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-hbase-compat-2.1.6	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-hbase-compat-2.2.5	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-hbase-compat-2.3.0	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-hbase-compat-2.4.0	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-hbase-compat-2.4.1	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-pherf	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-queryserver	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix-queryserver-client	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix-queryserver-it	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix-queryserver-load-balancer	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix-queryserver-orchestrator	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix-server-hbase-2.4	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix-tracing-webapp	5.1.1.7.2.16.0-287
	org.apache.phoenix	phoenix5-hive	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix5-hive-shaded	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix5-spark	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix5-spark-shaded	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix5-spark3	6.0.0.7.2.16.0-287
	org.apache.phoenix	phoenix5-spark3-shaded	6.0.0.7.2.16.0-287
	org.apache.phoenix.thirdparty	phoenix-shaded-commons-cli	1.1.0.7.2.16.0-287
	org.apache.phoenix.thirdparty	phoenix-shaded-guava	1.1.0.7.2.16.0-287
Apache Ranger	org.apache.ranger	conditions-enrichers	2.3.0.7.2.16.0-287
	org.apache.ranger	credentialbuilder	2.3.0.7.2.16.0-287
	org.apache.ranger	embeddedwebserver	2.3.0.7.2.16.0-287
	org.apache.ranger	jisql	2.3.0.7.2.16.0-287
	org.apache.ranger	ldapconfigcheck	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-adls-plugin	2.3.0.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.ranger	ranger-atlas-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-atlas-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-authn	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-distro	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-examples-distro	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-hbase-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-hbase-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-hdfs-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-hdfs-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-hive-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-hive-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-intg	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kafka-connect-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kafka-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kafka-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kms	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kms-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kms-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-knox-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-knox-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kudu-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kylin-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-kylin-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-metrics	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-nifi-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-nifi-registry-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-ozone-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-ozone-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-plugin-classloader	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-plugins-audit	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-plugins-common	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-plugins-cred	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-plugins-installer	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-policymigration	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-raz-adls	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-raz-chained-plugins	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-raz-hook-abfs	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-raz-hook-s3	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-raz-intg	2.3.0.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.ranger	ranger-raz-processor	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-raz-s3	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-raz-s3-lib	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-rms-common	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-rms-hive	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-rms-plugins-common	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-rms-webapp	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-s3-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-sampleapp-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-schema-registry-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-solr-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-solr-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-sqoop-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-sqoop-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-storm-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-storm-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-tagsync	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-tools	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-util	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-yarn-plugin	2.3.0.7.2.16.0-287
	org.apache.ranger	ranger-yarn-plugin-shim	2.3.0.7.2.16.0-287
	org.apache.ranger	sample-client	2.3.0.7.2.16.0-287
	org.apache.ranger	sampleapp	2.3.0.7.2.16.0-287
	org.apache.ranger	shaded-raz-hook-abfs	2.3.0.7.2.16.0-287
	org.apache.ranger	shaded-raz-hook-s3	2.3.0.7.2.16.0-287
	org.apache.ranger	ugsync-util	2.3.0.7.2.16.0-287
	org.apache.ranger	unixauthclient	2.3.0.7.2.16.0-287
	org.apache.ranger	unixauthservice	2.3.0.7.2.16.0-287
	org.apache.ranger	unixusersync	2.3.0.7.2.16.0-287
Apache Solr	org.apache.solr	solr-analysis-extras	8.4.1.7.2.16.0-287
	org.apache.solr	solr-analytics	8.4.1.7.2.16.0-287
	org.apache.solr	solr-cell	8.4.1.7.2.16.0-287
	org.apache.solr	solr-clustering	8.4.1.7.2.16.0-287
	org.apache.solr	solr-core	8.4.1.7.2.16.0-287
	org.apache.solr	solr-dataimporthandler	8.4.1.7.2.16.0-287
	org.apache.solr	solr-dataimporthandler-extras	8.4.1.7.2.16.0-287
	org.apache.solr	solr-jaegertracer-configurator	8.4.1.7.2.16.0-287
	org.apache.solr	solr-langid	8.4.1.7.2.16.0-287
	org.apache.solr	solr-ltr	8.4.1.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.solr	solr-prometheus-exporter	8.4.1.7.2.16.0-287
	org.apache.solr	solr-security-util	8.4.1.7.2.16.0-287
	org.apache.solr	solr-solrj	8.4.1.7.2.16.0-287
	org.apache.solr	solr-test-framework	8.4.1.7.2.16.0-287
	org.apache.solr	solr-velocity	8.4.1.7.2.16.0-287
Apache Spark	org.apache.spark	spark-avro_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-avro_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-catalyst_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-catalyst_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-core_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-core_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-graphx_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-graphx_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-hadoop-cloud_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-hadoop-cloud_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-hive_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-hive_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-kubernetes_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-kubernetes_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-kvstore_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-kvstore_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-launcher_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-launcher_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-mllib-local_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-mllib-local_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-mllib_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-mllib_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-network-common_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-network-common_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-network-shuffle_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-network-shuffle_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-network-yarn_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-network-yarn_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-repl_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-repl_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-shaded-raz	3.3.0.7.2.16.0-287
	org.apache.spark	spark-sketch_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-sketch_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-sql-kafka-0-10_2.11	2.4.8.7.2.16.0-287

Project	groupId	artifactId	version
	org.apache.spark	spark-sql-kafka-0-10_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-sql_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-sql_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-streaming-kafka-0-10_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-streaming-kafka-0-10_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-streaming_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-streaming_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-tags_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-tags_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-token-provider-kafka-0-10_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-token-provider-kafka-0-10_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-unsafe_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-unsafe_2.12	3.3.0.7.2.16.0-287
	org.apache.spark	spark-yarn_2.11	2.4.8.7.2.16.0-287
	org.apache.spark	spark-yarn_2.12	3.3.0.7.2.16.0-287
Apache Sqoop	org.apache.sqoop	sqoop	1.4.7.7.2.16.0-287
	org.apache.sqoop	sqoop-test	1.4.7.7.2.16.0-287
Apache Tez	org.apache.tez	hadoop-shim	0.9.1.7.2.16.0-287
	org.apache.tez	hadoop-shim-2.8	0.9.1.7.2.16.0-287
	org.apache.tez	tez-api	0.9.1.7.2.16.0-287
	org.apache.tez	tez-aux-services	0.9.1.7.2.16.0-287
	org.apache.tez	tez-common	0.9.1.7.2.16.0-287
	org.apache.tez	tez-dag	0.9.1.7.2.16.0-287
	org.apache.tez	tez-examples	0.9.1.7.2.16.0-287
	org.apache.tez	tez-ext-service-tests	0.9.1.7.2.16.0-287
	org.apache.tez	tez-history-parser	0.9.1.7.2.16.0-287
	org.apache.tez	tez-javadoc-tools	0.9.1.7.2.16.0-287
	org.apache.tez	tez-job-analyzer	0.9.1.7.2.16.0-287
	org.apache.tez	tez-mapreduce	0.9.1.7.2.16.0-287
	org.apache.tez	tez-protobuf-history-plugin	0.9.1.7.2.16.0-287
	org.apache.tez	tez-runtime-internals	0.9.1.7.2.16.0-287
	org.apache.tez	tez-runtime-library	0.9.1.7.2.16.0-287
	org.apache.tez	tez-tests	0.9.1.7.2.16.0-287
	org.apache.tez	tez-yarn-timeline-cache-plugin	0.9.1.7.2.16.0-287
	org.apache.tez	tez-yarn-timeline-history	0.9.1.7.2.16.0-287
	org.apache.tez	tez-yarn-timeline-history-with-acls	0.9.1.7.2.16.0-287
	org.apache.tez	tez-yarn-timeline-history-with-fs	0.9.1.7.2.16.0-287

Project	groupId	artifactId	version
Apache Zeppelin	org.apache.zeppelin	zeppelin-angular	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-display	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-interpreter	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-jdbc	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-jupyter	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-livy	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-markdown	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-server	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-shaded-raz	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-shell	0.8.2.7.2.16.0-287
	org.apache.zeppelin	zeppelin-zengine	0.8.2.7.2.16.0-287
Apache ZooKeeper	org.apache.zookeeper	zookeeper	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-client-c	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-contrib-loggraph	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-contrib-rest	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-docs	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-jute	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-recipes-election	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-recipes-lock	3.5.5.7.2.16.0-287
	org.apache.zookeeper	zookeeper-recipes-queue	3.5.5.7.2.16.0-287

What's New In Cloudera Runtime 7.2.16

You must be aware of the additional functionalities and improvements to features of components in Cloudera Runtime 7.2.16. Learn how the new features and improvements benefit you.

Support for Iceberg

Spark 3 now supports Iceberg.

What's New in Apache Atlas

Learn about the new features of Apache Atlas in Cloudera Runtime 7.2.16.

Relationship search

Entities in Atlas can be searched based on the relationships that describe various metadata between a couple of entity end-points.

See Relationship search for more information.

Basic search enhancement

While performing basic search operation in Atlas, you can exclude header attributes of entities from the response.

See Basic search enhancement for more information.

HDFS lineage extraction

Atlas supports the HDFS lineage data extraction mechanism.

See HDFS lineage data extraction for more information.

Viewing parent object for assigned classification or term in Atlas

Atlas supports viewing / searching for entities by using the assigned name for the classification / term.

See Parent object for assigned classification or term for more information.

Lineage on-demand

The on-demand lineage provides enhanced end user experience to handle data flow and related entities.

See On-demand lineage for more information.

Performance and Function Improvements

- Text-editor for Atlas parameters: While creating Classification, Glossary, and Business metadata, a new text editor is available. See Atlas Text-editor for more information.

What's New in Cruise Control

Learn about the new features of Cruise Control in Cloudera Runtime 7.2.16.

Rebasing Cruise Control to 2.5.85

Cruise Control is rebased to 2.5.85 version to be compatible with Apache Kafka 3.x version.

Adding MultiLevelRackAwareGoal to Cruise Control

As the currently available RackAwareGoal in Cruise Control does not support multi level rack awareness, a new goal was created to ensure that the replicas are assigned respecting the rules of multiple level racks.

The new goal is named com.cloudera.kafka.cruisecontrol.analyzer.goals.MultiLevelRackAwareDistributionGoal.

For more information, see the Multi-level rack-aware distribution goal documentation.

What's new in Data Analytics Studio

Learn about what is new in Data Analytics Studio (DAS) in Cloudera Runtime 7.2.16.

DAS has been deprecated

DAS has been deprecated in Cloudera Data Hub (Cloudera Runtime 7.2.16 release) and will be removed from the CDP stack in future releases. Cloudera encourages you to use Hue to run Hive LLAP workloads. You can submit queries from DAS, but cannot view query history without enabling the DAS Event Processor. For more information, see Enabling the DAS Event Processor.

What's New in Apache HBase

Learn about the new features of HBase in Cloudera Runtime 7.2.16.

COD supports JWT-based authentication for HBase clients

COD now supports JWT (JSON Web Token)-based authentication that uses an unique identifier and is a standard way of securely transmitting signed information between two parties. To know more about configuring JWT-based authentication for your HBase client, see Configuring JWT authentication for HBase client.

COD supports HBase metrics in Prometheus format

COD now supports exporting HBase metrics in Prometheus format. You can access HBase metrics in Prometheus format through the HBase web interface. For more information, see Accessing HBase metrics in Prometheus format.

What's New in Apache Hive

Learn about the new features of Hive in Cloudera Runtime 7.2.16.

Simplified Hive Warehouse Connector configuration

Setting up HWC configurations that are required by Spark is now simplified. As a cluster administrator, you need to specify the required configurations in Cloudera Manager as a one-time activity and then enable HWC by setting the spark.cloudera.useHWC property to "true". The spark.cloudera.useHWC property can either be specified in the spar k-defaults.conf file or by using the –conf option in spark-shell or spark-submit. For more information, see Setting up HWC configurations.

Hive Warehouse Connector Secure access mode

Hive Warehouse Connector (HWC) introduces the secure access mode that offers fine-grained access control (FGAC) column masking and row filtering to secure managed (ACID); or external, Hive table data that you query from Spark. Secure access mode requires you to set up staging location in your cloud storage service, such as S3 or ADLS, to temporarily store Hive files that users need to read from Spark. For more information, see Introduction to HWC Secure access mode.

Enable caching for Hive Warehouse Connector Secure access mode

You can enable caching for the HWC secure access mode to have finer control over read queries and ensure that the content updated outside of a Spark session is considered during reads. For more information, see Enabling caching for secure access mode

Hive ACID compaction observability

Compaction observability is a notification and information system based on metrics about the health of the compaction process. You can use Cloudera Manager to view compaction health checks for the Hive Metastore and Hive on Tez services, view actions and advice related to configurations and thresholds, and use the Compaction tab from the Hive Metastore service to view compaction-related charts based on the collected metrics. For more information, see Compaction Observability in Cloudera Manager.

Support for Hive Hybrid Procedural SQL

You can run Hive Hybrid Procedural SQL (HPL/SQL) queries from a client by connecting to Hive over JDBC. HPL/SQL is an Apache open source procedural extension for SQL for Hive users. For more information, see HPL/SQL stored procedures.

Hive Warehouse Connector support for Spark 3

As part of this release, Hive Warehouse Connector (HWC) is certified to work with Spark 3. You can use the binaries that are available in the /opt/cloudera/parcels/CDH/lib/hwc_for_spark3/ directory and use HWC to securely access Apache Hive managed tables from Spark. For more information, see Introduction to HWC.

Using SYS table to monitor compactions, transactions, and locks

You can monitor the progress and filter for specific compaction, transaction, and transaction lock jobs by querying the COMPACTIONS, TRANSACTIONS, and LOCKS view within the SYS database. For details, see Monitoring compactions, Monitoring transactions, and Monitoring transaction locks.

Support table defaults at database-level

You can use the database property, defaultTableType=EXTERNAL or ACID to specify the default table type to be created using the CREATE TABLE statement. You can specify this property when creating the database or at a later point using the ALTER DATABASE statement. For more information, see Understanding CREATE TABLE behavior.

Support external-only tables at database-level

You can choose to configure a database to allow only external tables to be created and prevent the creation of ACID tables. While creating a database, you can set the database property, EXTERNAL_TABLES_ONLY=true to ensure that only external tables are created in the database. For more information, see Understanding CREATE TABLE behavior.

Partition filtering support for the MSCK REPAIR TABLE statement

The MSCK REPAIR TABLE statement is enhanced to support filtering of the partition columns using operators so that a larger subset of partitions can be recovered (added/removed) without triggering a full repair. For more information, see Partition refresh and configuration.

Mapping specific columns in the INSERT clause of the MERGE statement

The MERGE statement is enhanced to support mapping of specific columns in the INSERT clause of the query instead of passing values (including null) for columns in the target table that do not have any data to insert. The unspecified columns in the INSERT clause are either mapped to null or use default constraints, if any. For more information, see Merging data in Hive tables.

Hive date and time UDF enhancements

The following Hive date and time user-defined functions (UDFs) are enhanced to use the DateTimeFormatter class instead of the SimpleDateFormat class, which may affect how date and timestamp values are parsed.

- unix_timestamp(): The unix_timestamp() function is enhanced to use the DateTimeFormatter class for String format dates instead of the SimpleDateFormat class. For details, see HIVE-25458.
- from_unixtime(): The from_unixtime() function is now enhanced to consider leap seconds. For details, see HIVE-25403.
- date_format(): The date_format() function that previously returned the output in UTC time zone is enhanced to display the default user session time zone. For details, see HIVE-25093.
- cast(): The cast() function is enhanced to display NULL when an incorrect date or timestamp is casted. Prior to this enhancement, when an incorrect date was casted, the function returned a converted value. For example, cast ('2020-20-20' as date) resulted in '2021-08-20' instead of NULL.

  This is because the DateTimeFormatter class that is used to parse string into date or timestamp was set to Resolver Style.LENIENT. This is now updated to use ResolverStyle.STRICT and returns NULL when an invalid date or timestamp is casted. For details, see HIVE-25306.

Configuration option to URL encode special characters in hbase.column.mapping

As part of this release, a new Hive configuration option is introduced to URL encode special characters like '#' or '%' that are used in hbase.columns.mapping values. The characters have to be encoded because the values are used to form the URI for Ranger based authentication. For more information, see Using Hbase Hive integration.

HMS support for external databases using DataConnectors (Technical Preview)

This release introduces the ability to map databases that reside in an external datasource, into a local Hive Metastore (HMS). The external datasources can be of different types, such as MySQL, Postgres, Redshift, or other HMS instances. Currently, we have support for external tables using StorageHandlers like JDBCStorageHandler or HBaseStorageHandler, however, the mapping needs to be configured for each table and can be cumbersome for a database with large volumes of tables.

DataConnector is a HMS object that contains definition or configuration details (hostname, credentials, etc) of a remote data source that are required to connect to the data source. Using DataConnectors, you can map an entire database instead of individual tables. The metadata for these tables are not persisted in Hive and are mapped and built during runtime. For more information, see HIVE-24396.

Technical Preview: This is a technical preview feature and considered under development. Do not use this in your production systems. To share your feedback, contact Support by logging a case on our Cloudera Support Portal. Technical preview features are not guaranteed troubleshooting guidance and fixes.

What's New in Hue

Learn about the new features of Hue in Cloudera Runtime 7.2.16.

Ability to rerun queries from the Job Browser page

A new option called Re Execute has been added to the Job Browser Queries page. You can select a query you want to rerun and click Re Execute. It takes you to the query editor to enable you to make changes and submit the query. For more information, see Rerunning a query from the Job Browser page.

Hue Query Processor scan frequency decreased to 5 minutes

The Hue Query Processor scans the event processor pipeline to retrieve the Hive query history and query details and displays them on the **Job Browser** page. The scan frequency has been decreased from 2 milliseconds to 5 minutes to optimize resource utilization. As a result, you may notice a delay in viewing the query history and query details on the **Job Browser** page for queries that finish executing in less than 5 minutes. However, you can still view the query history from the **Query history** tab below the query editor. See Configuring the Hue Query Processor scan frequency.

Ability to enable and disable auto-creation of user home directories in S3 and ADLS

If you have enabled fine-grained authorization to access S3 or ADLS, then Hue is configured to automatically create user home directories, by default. When a new user logs into Hue, Hue automatically creates a home directory for that user on S3 or ADLS. You can disable the automatic creation of user home directories by setting the autocreate_user_dir flag to false in the Hue Service Advanced Configuration Snippet. For more information, see Disabling automatic creation of user home directories on S3/ABFS.

Query Processor API to force data cleanup

Hue Query Processor cleans up queries older than a set number of days as per the set schedule. However, to manually clean up queries on a need basis, you can use the Query Processor API. When you call this API, it also runs a VACUUM command on the Query Processor tables. All queries that were run before the epoch time are cleared. For more information, see Ways to clean up old queries from the Query Processor tables.

Hue uses the SHA-256 signing algorithm for SAML authentication

The SHA-1 signing algorithm is deprecated in most environments. Hue now uses a stronger, secure hash algorithm, SHA-256, for signature and digest methods when authenticating using SAML.

Update to the list of supported non-ASCII characters

Hue supports an additional set of non-ASCII characters. % is also supported for file and folder names on HDFS and object stores. See Supported non-ASCII and special characters in Hue for a complete list.

Hue supports Spark SQL using Apache Livy

Hue supports Spark 3 and Livy 3. You can now run Spark SQL queries from Hue. Apache Spark and Apache Livy services are installed on your CDP cluster when you add a Data Hub cluster using the Data Engineering cluster template. See Enabling Spark 3 engine in Hue.

Hue supports rolling restart

Hue service downtime is reduced from more than 30 minutes to approximately 80-90 seconds when you restart the CDP cluster in the rolling restart mode. When you restart only the Hue service, then Hue's non-worker roles, such as the load balancer, Kerberos ticket renewer, and Hue server restart one after the other. For information about the rolling restart options, see Options to restart the Hue service.

Hue scripts included in CDP

CDP now includes Hue scripts that you can use for cleaning up old data, setting default editors, changing the document owners, and so on. You no longer need to clone the scripts from Cloudera's GitHub repository. For more information, see Using Hue scripts.

Support for HiveServer2 (HS2) high availability

Hue can now handle HS2 failover using ZooKeeper without setting up a load balancer. You must configure the following setting in the Hue Advanced Configuration snippet:

```
[beeswax]
hive_discovery_hs2=true
hive_discovery_hiveserver2_znode=/hiveserver2
```

See Configuring Hue to handle HS2 failover.

No 64-character restriction on hostnames for Hue roles

Hue supports creating Hue role hostnames of more than 64 characters. There is no longer a restriction of 64 characters for the "BalancerMember Route" property.

Hue uses TLS 1.2 by default

Hue and Hue Load Balancer use TLS 1.2 or 1.3 by default. You no longer have to configure settings in Cloudera Manager to enforce TLS 1.2.

What's new in Apache Iceberg

Learn about the new features of Iceberg in Cloudera Runtime 7.2.16.

Apache Iceberg v1 tables support is available on a GA (general availability) status in Data Hub with Hive, Impala and Spark compute engines. Iceberg is a cloud-native, open table format for organizing petabyte-scale analytic datasets on a file system or object store. You can deploy Iceberg based applications across multiple clouds including AWS, Azure and Google Cloud. Write once, run anywhere, and move from cloud to cloud. From a client, connect to Hive

or Impala in your Data Hub cluster, and run SQL commands on Iceberg tables. AWS and Azure environments are supported. New capabilities of Apache Iceberg in CDP include In-place Table Migration, Table Rollback: Table Maintenance, ORC open file format, Materialized Views, and SDX integration (Atlas). For more information about using Iceberg, see "Using Iceberg".

What's New in Apache Impala

Learn about the new features of Impala in Cloudera Runtime 7.2.16.

UTF-8 mode support

Some Impala STRING types now support UTF-8 aware behavior to ensure consistent results for non-ASCII characters in the string in both Hive and Impala.

Asynchronous model for some DDL statements

This release adds a new query option ENABLE_ASYNC_DDL_EXECUTION that you can configure to execute any request from an Impala client to the Impala server asynchronously in different threads without blocking the RPC. Using this asynchronous model, you can get a query handle and poll for state and results to avoid Impala clients hanging indefinitely.

Impala-shell with Python 3

Since Python 2.7 has reached the end of life, impala-shell can now be used with Python 3 by installing the latest release from PyPI at https://pypi.org/project/impala-shell/4.2.0a1/ .

BYTES function support

Impala now supports the BYTES() function. This function returns the number of bytes contained in a byte string.

Consolidating the ranger audit logs for the same table

Impala now consolidates the Ranger audit log entries of column accesses granted by the same policy for columns in the same table, after all the requests for accessing an object are processed.

Resolving ORC columns by names

Before this release, Impala resolved ORC columns by index. In this release, a query option ORC_SCHEMA_RESOLUTION is added to support resolving ORC columns by names.

Retrieving the data file name

Impala now supports including a virtual column in a standard SELECT statement select INPUT__FILE__NAME from <tablename> to retrieve the name of the data file that stores the actual row in a table.

Zipping unnest on arrays from Views

As part of this release, you can use zipping unnest functionality on arrays from Views. Before this release, this zipping functionality worked for arrays only in Tables but did not support Views as a source. For more information about using this zipping unnest functionality, see Zipping unnest on arrays from Views.

Min/Max filtering in Impala

Using Parquet format, you can query to find the minimum or maximum value for a column within a partition, row group, page, or row.

Reading and writing Parquet bloom filters

Bloom filter is a performance optimization feature now available in Impala. This filter tells you, rapidly and memory-efficiently, whether the data you are looking for is present in a file.

Added support for thrift-0.16.0

What's New in Apache Kafka

Learn about the new features of Apache Kafka in Cloudera Runtime 7.2.16.

Rebase on Kafka 3.1.2

Kafka shipped with this version of Cloudera Runtime is based on Apache Kafka 3.1.2. For more information, see the following upstream resources:

Apache Kafka Notable Changes:

- 3.0.0
- 3.0.1
- 3.1.0
- 3.1.1

Apache Kafka Release Notes:

- 3.0.0
- 3.0.1
- 3.1.0
- 3.1.1
- 3.1.2

Multi-level rack awareness

The rack awareness capabilities of Kafka have been improved to support multi-level cluster topologies. As a result, brokers can now be configured to run in a multi-level rack-aware mode. If this mode is enabled, the brokers provide multi-level rack awareness guarantees. These guarantees ensure that topic partition replicas are spread evenly across all levels of the physical infrastructure. For example, in a two-level hierarchy with Data Centers on the top level and racks on the second level, brokers will evenly spread replicas among both available DCs and racks.

The new mode is compatible with follower fetching. If multi-level mode is enabled, a compatible replica selector class is automatically installed. This implementation enables consumers (if configured), to fetch Kafka messages from the replica that is closest to them in the multi-level hierarchy.

Additionally, when Cruise Control is deployed on the cluster, the standard rack-aware goals in Cruise Control's configuration are replaced with a multi-level rack-aware goal. This goal ensures that Cruise Control optimizations do not violate the multi-level rack awareness guarantees. This goal is currently downstream only, available exclusively in Cloudera distributed Cruise Control. For more information, see the following resources:

- Kafka rack awareness
- Configure Kafka rack awareness
- Setting capacity estimations and goals

Expose log directory total and usable space through the Kafka API

KAFKA-13958 is backported in Kafka shipped with this version of Runtime. As a result, the Kafka API now exposes metrics regarding the total and usable disk space of log directories. The information on log directory space is collected by SMM and is exposed on the SMM UI. Specifically, you can now view the current log size of topics as

well as the total log size and remaining storage space of brokers. For more information on how you can monitor log size metrics on the SMM UI, see Monitoring log size information.

Kafka now accepts OAuthtokens that do not contain a “sub” claim

KAFKA-13730 is backported in Kafka shipped with this version of Runtime. As a result, Kafka now accepts OAuth tokens that do not contain the "sub" claim. If you are using OAuth tokens that do not contain a “sub” claim, the JWT Principal Claim Name For OAuth2 Kafka service property must be configured. This property specifies the claim that contains the client’s principal. For more information on OAuth2 authentication in Kafka, see OAuth2 authentication.

New Kafka connect connectors

The following new Kafka connect connectors are introduced:

- HDFS Stateless Sink
- Influx DB Sink
- Debezium Db2 Source [Technical Preview]

For more information, see Connectors.

Syslog TCP Source connector 2.0.0.

The Syslog TCP Source Kafka Connect connector is updated to version 2.0.0. The following notable changes and improvements are made:

- Three new properties are added, these are as follows:
  - Max Batch Size

    This property controls the maximum number of messages to add to a single batch of messages. This is a required property. Its default value is 1.
  - Authorized Issuer DN Pattern and Authorized Subject DN Pattern

    These properties allow you to enable authorization for incoming TLS connections. Both properties accept regular expressions as a value. The configured regular expressions are applied against the Distinguished Names of incoming TLS connections. If the Distinguished Names do not match the pattern, the following message is logged and the messages do not get forwarded to Kafka.

    ```
    Error: authorization failure
    ```

    Both properties are optional and are set to .* by default.
- The Max Number of TCP Connections property is replaced by the Max Number of Worker Threads property.

  Similarly to Max Number of TCP Connections, Max Number of Worker Threads is also used to specify the number of TCP connections, but instead of exactly specifying the number of allowed connections, you now specify how many worker threads are reserved for TCP connections. Note that a single worker thread is capable of handling multiple connections. This is a required property. Its default value is 2.
- Existing version 1.0.0. connectors will continue to function, upgrading them, however, is not possible. If you want to use the new version of the connector, you must deploy a new instance of the connector.
- Deploying a version 1.0.0. instance of the connector is no longer possible.

AvroConverter support for KConnect logical types

The AvroConverter now converts between Connect and Avro temporal and decimal types.

Connect internal topic Ranger policy

A new Ranger policy, connect internal - topic, is generated by default on fresh installations. This policy allows the Kafka and SMM service principals to access Kafka Connect internal topics (connect-configs, connect-offsets, connect-status) and the secret management's storage topic (connect-secrets).

Connector configurations must by default override the sasl.jaas.config property of the Kafka clients used by the connector

The Require Connectors To Override Kafka Client JAAS Configuration Kafka Connect property is now selected by default. This means that connector configurations must by default contain a sasl.jaas.config entry with an appropriate JAAS configuration that can be used to establish a connection with the Kafka service.

Connect JAAS enforcement now applies to non-override type Kafka client configs

When the Require Connectors To Override Kafka Client JAAS Configuration property is selected, the consumer.sasl. and producer.sasl. configurations are not emitted into the Connect worker configurations anymore. Additionally, the keytab name is randomized and the ${cm-agent:keytab} references in the Connector configurations will stop working.

What's New in Apache Kudu

Learn about the new features of Kudu in Cloudera Runtime 7.2.16.

Range-specific hash schemas for Kudu tables

Previously once the partition schema was set the number of hash buckets per range partition could not be changed. For the newly added range partitions the number of hash buckets was tied to the amount that was initialized in the partition schema. The custom hash schemas feature enables you to vary the number of hash buckets per range partition, both at table creation and alteration time.

For more information, see Managing Kudu tables with range-specific hash schemas.

Support native encryption at rest

Kudu now supports data encryption at rest, using File Key, Server Key and Cluster Key for encryption. However, data encryption at rest is supported only on fresh installation and once it is enabled you cannot disable it.

For more information, see Configuring data at rest encryption.

Support for Prometheus integration

Kudu now exposes metrics in Prometheus format for server level metrics. These metrics can be accessed using an endpoint at /metrics_prometheus path of the webserver for either Master or Tablet servers. Each metric is assigned a prefix master or tserver accordingly in order to differentiate between Master and Tablet servers running on the same host, along with sharing their unit type as a Prometheus label. For more information, see KUDU-3375.

Support range-aware rebalancing in Kudu CLI

The kudu cluster rebalance CLI tool can be used to run range-aware rebalancing.The range-aware rebalancing runs on a per-table basis, it can be run on one table at a time. To enable range-aware rebalancing for a particular table, you need to add the following two flags while invoking the tool:

```
--enable_range_rebalancing
--tables=<table_name_for_range_aware_rebalancing>
```

To perform cluster-wide rebalancing, it is recommended to run the ‘kudu cluster rebalance tool’. You can run the range-aware rebalancing, in addition to the cluster-wide rebalancing, for larger tables in the cluster that are multilevel-partitioned and can suffer from the hot-spotting issue.

For more information, see Run the tablet rebalancing tool.

Improvements

- KUDU-2181: Adding a new Kudu master instance to your cluster, for example to migrate to a multiple master configuration, is automated and does not require you to restart the already existing masters.

- KUDU-2623: Making KuduWriteOperation::table() method public to enable identification of the problematic table when an error happens.
- KUDU-3341: This patch improves catalog_manager's behavior when delete tablet with a 'WRONG_SERVER_UUID' error. This RetryTask is marked failed instead of getting retried in order to avoid too many requests.
- KUDU-3351: Add insert error count metrics in WriteResponsePB
- KUDU-3365: Expose INSERT/UPDATE metrics in the Java client API
- KUDU-3379: Make 'kudu table describe' output column comments
- KUDU-3389: support turning on/off auto rebalancer at runtime

What's New in Apache Livy

Learn about the new features of Livy in Cloudera Runtime 7.2.16.

High Availability support added for Livy

Livy now supports high availability. If there are more than one Livy Server in the cluster, high availability is automatically enabled.

What's New in Apache Ranger

The following new features and enhancements are generally available for Ranger customers in Cloudera Runtime 7.2.16:

Ranger Metrics on Audit Throughput

Ranger now provides a feature to monitor the throughput of audits generated by each plugin. Audit throughput monitoring includes an alert mechanism, which triggers when a large number of audits are generated and spool files are created. The Ranger UI now displays the audit metrics graphically. For more information, see the updated examples in Viewing Audit Metrics and Viewing Audit Details

New Ranger API to collect metrics in Ranger Admin

Ranger now provides two APIs to fetch ranger admin metrics One returns a response in JSON format and the other returns a response in prometheus-compatible format. For more information, see Ranger Admin Metrics API.

New Ranger API to collect metrics in Ranger RAZ

Ranger now provides two APIs to fetch ranger admin metrics One returns a response in JSON format and the other returns a response in prometheus-compatible format. For more information, see Ranger RAZ Metrics API.

Changes to Show Role Grant behavior

The Hive2 command line interface, Beeline returns role grant definitions for a specific principal, such as a user, group or role. For more information, see Showing Role|Grant definitions from Ranger HiveAuthorizer.

Changes to show Sync Source in Ranger User Management UI

The source type and details from which external users sync with Ranger now appears in Ranger Admin UI. For more information, see the updated examples throughout: Administering Ranger Users, Groups, Roles, and Permissions.

Provided Ranger support on DataHub HDFS

Spark jobs now interact with HDFS for scratch/staging data that cannot rely on S3. Reduced dramatic performance degradation due to lack of atomic rename. Added to 7.2.14+ via hotfix. For more information, see the updated examples in Enabling Ranger HDFS plugin manually on a Data Hub .

Performance and Function Improvements

- Performance improvements for Ranger log file rotation implemented and documented. For 2more information, see Managing logging properties for Ranger services.
- Prior to 7.1.8 Ranger RMS download mapping API was an open API. Now it is made secured and it requires JWT/Kerberos authentication to access this API.
- New Audit filters in HDFS audits are added to exclude the ACID operations which are flooding the ranger audits and are not really needed.
- Consolidate policies created for {OWNER} by Authzmigrator. Provided a feature to skip the {OWNER} policy in authzmigration tool.

What's New in Schema Registry

Learn about the new features of Schema Registry in Cloudera Runtime 7.2.16.

Schema Registry instances behind load balancer

You can now use load balancer in front of Schema Registry instances. It is very common to have multiple instances of the same application and have a load balancer in front of them. This can be useful for failover reasons in HA environments, and it can also help sharing the load between instances. You can also use load balancer in front of Schema Registry instances in an environment with Kerberos or SSL enabled.

AvroConverter support for KConnect logical types

AvroConverter now converts between Connect and Avro temporal and decimal types.

Support for alternative jersey connectors in SchemaRegistryClient

connector.provider.class can be configured in Schema Registry Client. If it is configured, schema.registry.client.retry.policy should also be configured to be different than default.

This also fixes the issue with some 3rd party load balancers where the client is expected to follow redirects and authenticate while doing that.

Retrieve principal from client's certificate

When two-way TLS authentication is enabled, Schema Registry extracts the principal from the certificate and uses it for authentication or authorization.

Schema Registry CDC support - change default schema compatibility

When a new Avro schema is created and its compatibility is not explicitly set, then a default compatibility value is used. Until now, that value was always BACKWARD. After this change, users on the server side can configure the default value.

Schema Registry with Knox uses round-robin load balancing

When multiple instances of Schema Registry are running, Knox uses round-robin to forward the requests.

Upgraded Avro version to 1.11.1

Avro got upgraded from version 1.9.1 to 1.11.1.

What's New in Apache Spark

Learn about the new features of Spark in Cloudera Runtime 7.2.16.

Support for Iceberg

Spark 3 now supports Iceberg.

Updated Spark 3 version of CDP

Spark 3 version of CDP stack was updated to 3.3.

Apache Spark 3 version support

- Support for virtual clusters powered by Apache Spark 3 is now available.
- The following functionalities are not currently supported:
  - Deep analysis (visual profiler)
  - HWC - that is, Hive managed ACID tables (Direct Reader & JDBC mode)
  - Phoenix Connector
  - SparkR

See Running Apache Spark 3 applications and Data Engineering clusters.

What's New in Sqoop

Learn what's new in the Apache Sqoop client in Cloudera Runtime 7.2.16.

To access the latest Sqoop documentation on Cloudera's documention web site, go to Sqoop Documentation 1.4.7.7.1.6.0.

Discontinued maintenance of direct mode

The Sqoop direct mode feature is no longer maintained. This feature was primarily designed to import data from an abandoned database, which is no longer updated. Using direct mode has several drawbacks:

- Imports can cause an intermittent and overlapping input split.
- Imports can generate duplicate data.
- Many problems, such as intermittent failures, can occur.
- Additional configuration is required.

Do not use the --direct option in Sqoop import or export commands.

What's new in Streams Messaging Manager

Learn about the new features of Streams Messaging Manager in Cloudera Runtime 7.2.16.

Improved alter topic functionalities

You can now increase the number of partitions of a topic (but not decrease). The option is available on the Configs tab on the Topic Details page.

Partition Assignment tab on the Topic Details page

The Assignment tab, on the topic details page, shows the current state of the partitions and replicas of the topic. It shows some topic-level statistics and the replica assignment of all partitions. If rack awareness is being used in the Kafka cluster, the replica assignment is shown in a rack-based view. If the rack IDs follow the format of multi-level rack IDs, the rack IDs are rendered as a hierarchy.

Added sorter functionality to partition lists in SMM UI

The SMM UI contains Brokers and Topics pages where records contain broker or topic specific partition lists and their profile pages as well. All partition list columns become sortable.

SMM UI shows broker rack information

The Brokers page and the Broker Details page now both show the rack ID of the brokers. SMM now also supports a new endpoint: /api/v1/admin/topics/{topicName}/description.

Improved SMM UX for Kafka Connectors configuration

The connector selection and connector configuration workflow steps are now separated into two different steps. Search and autocomplete are now available for Connect configuration keys. The help icon provides detailed information about each configuration key. The data type of the configuration values can be chosen from the options menu. For more information, see *Setting connector configurations*.

Added partition log-size information to the SMM UI

The SMM UI shows log-size related information about brokers, topics, and partitions. Furthermore, warning messages appear when log directory related errors are reported by Kafka.

SMM connector profile shows connector level error

Errors causing the whole connector to fail are now displayed on the connector profile page if available.

Data explorer allows specifying consumer isolation.level

"/api/v1/admin/topics/{topicName}/partition/{partitionId}/payloads" endpoint has a new parameter: "consumerIsolationLevel". The accepted values are "read_committed" and "read_uncommitted". This sets the "isolation.level" config for the KafkaConsumer used for retrieving messages. The default value is "read_uncommitted".

Additionally, the parameter can be set on the UI as well.

The Data Explorer page has a new design

The Data Explorer now uses the “from offset” and “record limit” parameters to select an offset window to query.

Offset-lag metrics in SMM UI

SMM UI Replications tab now also shows the replication-records-lag metric.

SMM UI Data Explorer shows null values explicitly

Data Explorer in the SMM UI now displays 'null' with italic style applied when the value is null rather than an empty value.

SMM uses a specific REST API to fetch list of topics

SMM Connect page now uses the Connect active topic tracking feature to list the topics used by the Connectors instead of the connector configuration's topic property. Sink Connectors show up based on which topics they consumed from (regardless of whether "topics" or "topics.regex" config was used), and Source Connectors show up based on which topics they produced into.

Improved configuration of SMM Kafka interceptors

New configuration prefix for SMM monitoring interceptor's background producer: "smm.monitoring.interceptor.producer.".

Clients that use either of the SMM monitoring interceptors (MonitoringConsumerInterceptor, MonitoringProducerInterceptor) use a background producer to push client metrics into Kafka every 30 seconds. This background producer now can be configured by passing Producer Configurations to the client that uses the interceptor with the "smm.monitoring.interceptor.producer." prefix. The prefix is trimmed and the remaining part of the configuration is passed to the background producer. For instance, if the user wishes to configure the "batch.size" property for the background producer, the following configuration should be passed: "smm.monitoring.interceptor.producer.batch.size".

There is a different behavior, however, with the "client.id" property. If the user does not provide a configuration to the client id (smm.monitoring.interceptor.producer.client.id), the default is used, which is: "smm-monitoring-interceptor".

SMM Data Explorer shows JSON output in pretty printed format

SMM Data Explorer can show JSON output in pretty printed format by using the "JSON Pretty Print" deserializer.

What's New in Streams Replication Manager

Learn about the new features of Streams Replication Manager in Cloudera Runtime 7.2.16.

Internal SRM topics are now automatically added to the replication deny list

SRM now ignores internal SRM topics when replicating data. If required, this behavior can be disabled by adding srm.internal.topic.exclude.enable=false to Streams Replication Manager's Replication Configs in Cloudera Manager.

Metrics and health checks for the status processor Streams application in SRM Service

SRM Service health tests now show the state of the Connect status processor Streams application.

SRM topic creation timeout increased

Streams Replication Manager internal topic creation timeout property defaults are increased to 20s to tolerate intermittent issues at startup.

Improvements related to raw metric collection and aggregation

Replication status metrics are enhanced with task information

The SRM Service now reports the status of the connectors and tasks in the replication status with detailed descriptions. The status and detailed descriptions can be viewed on the Replications tab of SMM UI.

New metric replication-records-lag

The SRM Service now reports a new metric on its REST API called replication-records-lag. This metric provides information regarding the replication lag based on offsets. The metric is available both on the cluster and the topic level.

Raw metrics are compressed using LZ4

SRM internal metric producers from now on use LZ4 compression by default. LZ4 was chosen as it provides the best combination in terms of compression speed and performance. As a result, Cloudera recommends that you use LZ4. If required, however, you can change the compression by doing the following:

1. Add the following configuration entries to Streams Replication Manager's Replication Configs

```
workers.cloudera.metrics.reporter.producer.compression.t
ype=[***COMPRESSION***]
connectors.cloudera.metrics.reporter.producer.compression.t
ype=[***COMPRESSION***]
```

2. Add the following to Additional Configs For Streams Application Running Inside SRM Service

```
producer.compression.type=[***COMPRESSION***]
```

For more information regarding, metrics, monitoring, as well as raw metric collection and aggregation, see Streams Replication Manager monitoring and metrics .

What's New in Apache Hadoop YARN and YARN Queue Manager

Learn about the new features of Hadoop YARN and YARN Queue Manager in Cloudera Runtime 7.2.16.

Apache Hadoop YARN

There are no new features for Apache Hadoop YARN in this release of Cloudera Runtime.

YARN Queue Manager

Dynamic Queue Scheduling

Dynamic Queue Scheduling is now generally available and can be used in production environments. This is the result of multiple changes, improvements, and new features such as Dynamic Configuration revalidation and execution logs.

For more information, see Dynamic Queue Scheduling.

Queue priority

Setting queue priorities is now supported by the YARN Queue Manager UI. By setting queue priorities you can ensure that applications can access cluster resources. This is especially important in the case of Hive LLAP, long-running applications, and applications that require large containers.

For more information, see Setting queue priorities.

Setting Maximum Parallel Application Limits

You can set the maximum number of applications limits for all queues, all users, and at the user level. The maximum parallel application limit is inherited from the “root” queue level and is lowered down in the queue hierarchy. The limit is checked in the queue hierarchy and the lowest value is applied as the limit.

For more information, see Setting Maximum Parallel Application.

Editing placement rules

Support to edit previously created placement rules was added.

For more information, see Editing placement rules.

Refresh queues option in Queue Manager UI

A Refresh button was added to the Overview tab in the YARN Queue Manager UI which provides the functionality to refresh the queues on demand.

Configuring the the capacity and max capacity of root queue in absolute mode

Support to configure memory/vcores and maximum memory/vcores for the root queue in absolute resource allocation mode is added. They can be set using the YARN Queue Manager UI.

For more information, see Configuring the resource capacity of root queue in absolute mode.

New YARN Queue Manager Overview Page

The new YARN Queue Manager **Overview** page has a new improved User Interface (UI) with the following new features:

- Minimap: The Overview page now has a minimap of the queue structure. It shows the whole queue structure even if you zoom in to a specific part of it.
- Refresh: You can click the Refresh icon for in-screen refresh of the page.
- Zoom and Panning : You can use the mouse to zoom in and zoom out on the screen to view the queue structure. You can also drag the queue structure to see different parts of the structure.
- Tool Tip: You can hover on the queue name for information like queue name and its queue path, queue status, and capacity. Previously, only the queue name and its path was displayed.

Unaffected Components in this release

There are no new features for the following components in Cloudera Runtime 7.2.16.

- Data Analytics Studio
- Apache Hadoop HDFS
- Apache Knox

- Apache Oozie
- Apache Ozone
- Apache Phoenix
- Apache Solr
- Apache ZooKeeper

Fixed Issues In Cloudera Runtime 7.2.16

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.

Fixed Issues in Atlas

Review the list of Apache Atlas issues that are resolved in Cloudera Runtime 7.2.16.
CDPD-46492: Relationship Attribute Filter button was stuck in loading state for specific internal relationships because of the missing dependencies. The dependencies has been added and the issue is resolved now.

CDPD-43821: Text-editor opens now for even when attribute key of Business metadata is space separated.

CDPD-42937: The broken image issue for expand button was due to the licence added in SVG file and now is fixed and the expand button is rendered on UI.

CDPD-42916: Made the required changes for commons-codec version in XML file.

CDPD-42913: Opensearch Support for Ranger access audits.

CDPD-42633: ElasticSearch dependency is removed from Atlas.

CDPD-40822: Upgrade Spring Framework to 5.3.20 due to CVE-2022-22971, CVE-2022-22970.

CDPD-40707: Parents name for assigned Classification & Glossary name for assigned terms was not seen on detail page and search result page, by providing this fix we are able to see the parent name on mouse hover of assigned terms and classifications. This is resolved now.

CDPD-36992: This changes implements a check on relationship edges, identifying already deleted relationship edges and avoiding invocation of delete method on them.

Thus only allowing deletion of active relationship edges.

CDPD-35953: Imapala process entities created by ImpalaHook saves query-string in name field. Since query-string can be large, we are getting the longer than the max error.

To store qualifiedName in name field instead of query-string

CDPD-35212: If Kerberos is enabled, - kinit -kt /cdep/keytabs/atlas.keytab atlas@ROOT.HWX.SITE - Add below to DEFAULT_JVM_OPTS in repair_index.py - Djavax.security.auth.useSubjectCredsOnly=false -Djava.security.auth.login.config=<path to atlas_jaas.conf> If SSL is enabled, we need to make sure Solr cert or RootCA certificate is added which make use of below atlas-application properties - keystore.file <path to keystore jks file> - truststore.file <path to truststore jks file> - cert.stores.credential.provider.path <path to jceks file> export HADOOP_CREDSTORE_PASSWORD=<password>

CDPD-30950: Instead of loading entire lineage, fetch lineage on demand to improve performance.

CDPD-28569: Added support for searching relationships in Atlas.

OPSAPS-57415: "This fix applies to CDP Private Cloud Base, Public Cloud and Private Cloud Data Services Required Kafka topics for Atlas will be pre-created for fresh install on clusters for kerberos enabled environments."

OPSAPS-62184: HDFS Lineage configurations are available and user can configure the Blacklist and Whitelist paths from Atlas configurations.

OPSAPS-63767: Hadoop Group based authorization should work for Atlas Ranger plugin.

OPSAPS-64201: Default value set false for Atlas Server / Gateway Roles and Hook configurations for both 7.2.16 and 7.1.8 CSD.'s

CDPD-48122: Operations like admin/audits, admin/purge fail with a 500 internal server error message

Before the update, the attributes of __AtlasAuditEntry type where not getting indexed in the Solr, because the entities of __AtlasAuditEntry gets created before the attributes gets indexed (TypeDef gets created).

Apache patch information

- ATLAS-4710
- ATLAS-4673
- ATLAS-4572
- ATLAS-4442
- ATLAS-4610
- ATLAS-4663
- ATLAS-4558
- ATLAS-4440
- ATLAS-4571
- ATLAS-4641

Fixed Issues in Avro

There are no fixed issues for Avro in Cloudera Runtime 7.2.16.

Apache patch information

None

Fixed Issues in Cloud Connectors

Review the list of Cloud Connectors issues that are resolved in Cloudera Runtime 7.2.16.
CDPD-43464: The aws-java-sdk library was updated to 1.12.262+ due to CVE-2022-31159. Note that the s3a connector has never been vulnerable to the CVE, as it does not use the SDK's TransferManager for downloading files.

Fixed issues in Cruise Control

There are no fixed issues for Cruise Control in Cloudera Runtime 7.2.16.

Fixed issues in Data Analytics Studio

Review the list of Data Analytics Studio (DAS) issues that are resolved in Cloudera Runtime 7.2.16.
OPSAPS-64287: DAS WebUI fails to open with the "Request Header Fields Too Large" error

This issue has been fixed by adding a new optional parameter called das_application_connector_configs to configure the header size.

Fixed Issues in Apache Hadoop

Review the list of Hadoop issues that are resolved in Cloudera Runtime 7.2.16.
CDPD-40841: HADOOP-17844 - Upgrade json-smart to 2.4.7 due to CVE-2021-31684.

Apache Patch Information

- HADOOP-17844
- HADOOP-18120

Fixed Issues in HBase

Review the list of HBase issues that are resolved in Cloudera Runtime 7.2.16.
CDPD-46164: JWT related JARS are added to the HBASE Tarball.

CDPD-45126: To configure MCC as a drop-in replacement when MCC jar is part of the classpath, you need the following configurations

- * hbase.client.connection.impl = com.cloudera.hbase.mcc.ConnectionMultiCluster
- * hbase.mcc.create.token.manager = false
- * hbase.mcc.token.file.name = ""

CDPD-44124: Missing dependency jar file is added to the HBaseRangerPlugin.

OPSAPS-64485: Added support for HBase native TLS implementation. When Auto-TLS is enabled, the feature is enabled by default; however with plaintext fallback support to ensure backward compatibility.

OPSAPS-64940: The issues related to HBase backup or restore API in GOV clouds are fixed. The APIs were not working because some safety valves were not generated.

Apache Patch Information

Bugfixes:

- HBASE-27484 FNFE on StoreFileScanner after a flush followed by a compaction
- HBASE-27407 Fixing check for "description" request param in JMXJsonServlet.java (#4816)
- HBASE-27362 CompactSplit.requestCompactionInternal may bypass compactionsEnabled check (#4768)
- HBASE-22939 SpaceQuotas - Bulkload from different hdfs failed when space quotas are turned on. (#4750)
- HBASE-27246 RSGroupMappingScript#getRSGroup has thread safety problem (#4657)
- HBASE-27268 In trace log mode, the client does not print callId/startTime and the server does not print receiveTime (#4710)
- HBASE-23330: Fix delegation token fetch with MasterRegistry (#1084) (#4598)
- HBASE-27292. Fix build failure against Hadoop 3.3.4 due to added dependency on okhttp. (#4687)
- HBASE-27244 bin/hbase still use slf4j-log4j while reload4j in place (#4652)
- HBASE-27232 Fix checking for encoded block size when deciding if bloc… (#4640)
- HBASE-27282 CME in AuthManager causes region server crash (#4684)
- HBASE-27275 graceful_stop.sh unable to restore the balance state (#4680)
- HBASE-27053 IOException during caching of uncompressed block to the block cache (#4610)
- HBASE-27097 SimpleRpcServer is broken (#4613)
- Amend HBASE-27180 Fix multiple possible buffer leaks (#4597)
- HBASE-27189 NettyServerRpcConnection is not properly closed when the netty channel is closed (#4611)
- HBASE-27180 Fix multiple possible buffer leaks (#4597)

- HBASE-26708 Netty leak detected and OutOfDirectMemoryError due to direct memory buffering with SASL implementation (#4596)
- HBASE-27175 - Failure to cleanup WAL split dir log should be at INFO level (#4593)
- HBASE-27171 Fix Annotation Error in HRegionFileSystem (#4588)
- HBASE-23330: Fix delegation token fetch with MasterRegistry (#1084) (#4598)
- HBASE-27170 ByteBuffAllocator leak when decompressing blocks near minSizeForReservoirUse (#4592)
- HBASE-26790 getAllRegionLocations can cache locations with null hostname (#4575)
- HBASE-26945 Quotas causes too much load on meta for large clusters (#4576)
- HBASE-26856 BufferedDataBlockEncoder.OnheapDecodedCell value can get corrupted
- Revert "HBASE-25709 Close region may stuck when region is compacting and skipped most cells read (#3117)" (#4524)
- HBASE-27097 SimpleRpcServer is broken (#4521)
- HBASE-26985 check permission for SecureBulkLoadManager (#4379)
- HBASE-27046 The filenum in AbstractFSWAL should be monotone increasing (#4449)
- HBASE-26680 Close and do not write trailer for the broken WAL writer(addendum) (#4405)
- HBASE-26523 + HBASE-25465 + HBASE-26855 backport to branch-2.4 (#4439)
- HBASE-27017: MOB snapshot is broken when FileBased SFT is used (#4466)
- HBASE-27069 Hbase SecureBulkload permission regression (#4475)
- HBASE-27061 two phase bulkload is broken when SFT is in use. (#4465)
- HBASE-27024 The User API and Developer API links are broken on hbase.apache.org (#4424)
- HBASE-27032 The draining region servers metric description is incorrect (#4428)
- HBASE-25058 Export necessary modules when running under JDK11
- HBASE-27033 Backport "HBASE-27013 Introduce read all bytes when using pread for prefetch" (#4429)
- HBASE-27021 StoreFileInfo should set its initialPath in a consistent way (#4419)
- HBASE-26994 MasterFileSystem create directory without permission check (#4391)
- HBASE-26963 ReplicationSource#removePeer hangs if we try to remove bad peer. (#4413)
- HBASE-26971 SnapshotInfo --snapshot param is marked as required even when trying to list all snapshots (#4366)
- HBASE-26941 LocalHBaseCluster.waitOnRegionServer should not call join while interrupted (#4352)
- HBASE-26944 Possible resource leak while creating new region scanner (#4339)
- HBASE-26938 Compaction failures after StoreFileTracker integration (#4350)
- HBASE-26895 on hbase shell, 'delete/deleteall' for a columnfamily is not working (#4283)
- HBASE-26901 delete with null columnQualifier occurs NullPointerException when NewVersionBehavior is on (#4295)
- HBASE-26920 Fix missing braces warnings in TestProcedureMember (#4315)
- HBASE-26916 Fix missing braces warnings in DefaultVisibilityExpressionResolver (#4313)
- HBASE-26812 ShortCircuitingClusterConnection fails to close RegionScanners when making short-circuited calls (#4302)

Improvements:

- HBASE-27406 Make /prometheus endpoint accessible from HBase UI (#4833)
- HBASE-27224 HFile tool statistic sampling produces misleading results (#4638)
- HBASE-27229 BucketCache statistics should not count evictions by hfile (#4639)
- HBASE-27395 Adding description to Prometheus metrics (#4807)
- HBASE-27381 Still seeing 'Stuck' in static initialization creating RegionInfo instance (#4813)
- HBASE-27403 Remove 'Remove unhelpful javadoc stubs' spotless rule for now (#4809)
- HBASE-27391 Downgrade ERROR log to DEBUG in ConnectionUtils.updateStats (#4804)
- HBASE-27384 Backport HBASE-27064 (CME in TestRegionNormalizerWorkQueue) to 2.4 (#4794) (#4468)
- HBASE-27393 Frequent and not useful "Final timeLimitDelta" log lines (#4802)
- HBASE-27365 Minimise block addition failures due to no space in bucket cache writers queue by introducing wait time

- HBASE-27386 Use encoded size for calculating compression ratio in block size predicator (#4795)
- HBASE-27370 Avoid decompressing blocks when reading from bucket cache… (#4781)
- HBASE-27373 Fix new spotbugs warnings after upgrading spotbugs to 4.7.2 (#4787) (#4791)
- HBASE-27368 Do not need to throw IllegalStateException when peer is not active in ReplicationSource.initialize (#4779)
- HBASE-27371 Bump spotbugs version (#4783)
- HBASE-27332 Remove RejectedExecutionHandler for long/short compaction thread pools (#4731)
- HBASE-27335 HBase shell hang for a minute when quiting (#4737)
- HBASE-27317 Rectifying the option for columnfamily as mandatory (#4773)
- HBASE-27313: Persist list of Hfiles names for which prefetch is done (#4771)
- HBASE-20904 Prometheus /metrics http endpoint for monitoring (#4691)
- HBASE-25922 - Disabled sanity checks ignored on snapshot restore (#4533) (#4734)
- HBASE-27320 hide some sensitive configuration information in the UI (#4723)
- HBASE-27089 Add “commons.crypto.stream.buffer.size” configuration (#4491)
- HBASE-27294 Add new hadoop releases in our hadoop checks (#4692)
- HBASE-27296 Some Cell's implementation of toString() such as IndividualBytesFieldCell prints out value and tags which is too verbose (#4695) (#4703)
- HBASE-27221 Bump spotless version to 2.24.1 (#4693)
- HBASE-27265 : Tool to read StoreFileTrackerFile (#4673)
- HBASE-27301 Add Delete addFamilyVersion timestamp verify (#4700)
- HBASE-27281 Add default implementation for Connection$getClusterId (#4683)
- HBASE-27264 Add options to consider compressed size when delimiting blocks during hfile writes (#4675)
- HBASE-27293 Remove jenkins and personality scripts support for 1.x (#4690)
- HBASE-26775 - add synchronized modifier to the toString() method of ProcedureEvent.java (#4681)
- HBASE-27273 Should stop autoRead and skip all the bytes when rpc request too big (#4679)
- HBASE-27269 The implementation of TestReplicationStatus.waitOnMetricsReport is incorrect (#4678)
- HBASE-27271 BufferCallBeforeInitHandler should ignore the flush request (#4676)
- HBASE-27257 Remove unnecessary usage of CachedBlocksByFile from RS UI (#4667)
- HBASE-27087 TestQuotaThrottle times out
- Revert "HBASE-23330: Fix delegation token fetch with MasterRegistry (#1084) (#4598)"
- HBASE-27225 Add BucketAllocator bucket size statistic logging (#4637)
- HBASE-27204 BlockingRpcClient will hang for 20 seconds when SASL is enabled after finishing negotiation (#4642)
- HBASE-27219 Change JONI encoding in RegexStringComparator (#4632)
- HBASE-27220 Apply the spotless format change in HBASE-27208 to our code base
- HBASE-27208 Use spotless to purge the missing summary warnings from error prone (#4628)
- HBASE-27205 Fix tests that rely on EnvironmentEdgeManager in branch-2.4 (addendum) (#4631)
- HBASE-27211 Data race in MonitoredTaskImpl could cause split wal failure (#4630)
- HBASE-27205 Fix tests that rely on EnvironmentEdgeManager in branch-2.4 (#4625)
- HBASE-27192 The retry number for TestSeparateClientZKCluster is too small (#4614)
- HBASE-27193 TestZooKeeper is flaky (#4615)
- HBASE-27161 Improve TestMultiRespectsLimits (#4586)
- HBASE-27188 Report maxStoreFileCount in jmx (#4609)
- HBASE-27186 Report block cache size metrics separately for L1 and L2 (#4608)
- HBASE-27048 Server side scanner time limit should account for time in queue (#4562)
- HBASE-27169 TestSeparateClientZKCluster is flaky (#4587)
- HBASE-27160 ClientZKSyncer.deleteDataForClientZkUntilSuccess should break from the loop when deletion is succeeded (#4579)
- HBASE-27060 Addendum spotless fix (#4580)
- HBASE-27060 Addendum fix HBaseTestingUtility import in test

- HBASE-27060 Allow sharing connections between AggregationClient instances (#4566)
- HBASE-27146 Avoid CellUtil.cloneRow in MetaCellComparator (#4571)
- HBASE-27151 TestMultiRespectsLimits.testBlockMultiLimits repeatable failure
- HBASE-27050 Support unit test pattern matching again (#4447)
- HBASE-27141 Upgrade hbase-thirdparty dependency to 4.1.1 (#4552)
- HBASE-27051 TestReplicationSource.testReplicationSourceInitializingMetric is flaky (#4448)
- HBASE-27143 Add hbase-unsafe as a dependency for a MR job triggered by hbase shell (#4554)
- HBASE-27099 The log printing fspread/fsread cost time unit should be milliseconds (#4500)
- HBASE-27128 when open archiveRetries totalLogSize calculation mistake (#4546)
- HBASE-27125 The batch size of cleaning expired mob files should have an upper bound(addendum) (#4553)
- HBASE-27125 The batch size of cleaning expired mob files should have an upper bound (#4541)
- HBASE-27117 Update the method comments for RegionServerAccounting (#4532)
- Revert "HBASE-27084 Add spotless:check in mvn verify stage (#4482)"
- HBASE-26923 PerformanceEvaluation support encryption option (#4489)
- HBASE-27095 HbckChore should produce a report
- HBASE-27093 AsyncNonMetaRegionLocator#put Complete CompletableFuture outside lock block (#4496)
- HBASE-27038 CellComparator should extend Serializable (#4492)
- HBASE-27084 Add spotless:check in mvn verify stage (#4482)
- HBASE-27080 Optimize debug output log of ConstantSizeRegionSplitPolicy class. (#4481)
- HBASE-26649 Support meta replica LoadBalance mode for RegionLocator#getAllRegionLocations() (#4442) (#4485)
- HBASE-27039 Some methods of MasterRegion should be annotated for testing only (#4433)
- HBASE-27023 Fix license issues after running spotless:apply (#4458)
- HBASE-27079 Lower some DEBUG level logs in ReplicationSourceWALReader to TRACE (#4476)
- HBASE-26933 Addendum remove unused resources and links on site
- HBASE-26933 Remove all ref guide stuff on branch other than master (#4426)
- HBASE-27030 Fix undefined local variable error in draining_servers.rb (#4427)
- HBASE-27047 Fix typo for metric drainingRegionServers (#4441)
- HBASE-27027 Use jetty SslContextFactory.Server instead of deprecated SslContextFactory (#4425)
- HBASE-27054 TestStochasticLoadBalancerRegionReplicaLargeCluster.testRegionReplicasOnLargeCluster is flaky (#4454)
- HBASE-27018 Add a tool command list_liveservers (#4416)
- HBASE-27006 Move nightly integration testing to new larger test node class. (#4438)
- HBASE-27045 Disable TestClusterScopeQuotaThrottle (#4440)
- HBASE-27003 Optimize log format for PerformanceEvaluation (#4411)
- HBASE-27000 Block cache stats (Misses Caching) display error in RS web UI (#4406)
- HBASE-26995 Remove ref guide check in pre commit and nightly for branches other than master (#4399)
- HBASE-26990 Add default implementation for BufferedMutator interface setters (#4387)
- HBASE-26892 Add spotless:check in our pre commit general check (#4393)
- HBASE-26899 Run spotless:apply
- HBASE-26617 Use spotless to reduce the pain on fixing checkstyle issues
- HBASE-26674 Should modify filesCompacting under storeWriteLock (#4040)
- HBASE-26999 HStore should try write WAL compaction marker before repl… (#4407)
- HBASE-26860 Backport " HBASE-25681 Add a switch for server/table queryMeter" to branch-2.4 (#4240)
- HBASE-26917 Do not add --threads when running 'mvn site' (#4354)
- HBASE-26932 Skip generating ref guide when running 'mvn site' on branch other than master (#4360)
- HBASE-26980 Update javadoc of BucketCache.java (#4374)
- HBASE-26581 Add metrics for failed replication edits (#4347)
- HBASE-26942 cache region locations when getAllRegionLocations (#4357)
- HBASE-24337 Backport HBASE-23968 to branch-2 (#3588)

- Revert "HBASE-25665 Option to use hostname instead of canonical hostname for secure HBase cluster connection (#3051)"
- HBASE-26618 Involving primary meta region in meta scan with CatalogRe… (#4321) (#4328)
- HBASE-26880 Misspelling commands in hbase shell will crash the shell (#4325)
- HBASE-26922 Fix LineLength warnings as much as possible if it can not be fixed by spotless (#4324)
- HBASE-26929 Upgrade surefire plugin to 3.0.0-M6 (#4319)
- HBASE-26928 Fix several indentation problems (#4323)
- HBASE-26882 Backport " HBASE-26810 Add dynamic configuration support f… (#4278)
- HBASE-26885 Addendum throw exception instead of return in TRSP to let the procedure retry (#4299)
- HBASE-26921 Rewrite the counting cells part in TestMultiVersions (#4316)
- HBASE-26919 Rewrite the counting rows part in TestFromClientSide4 (#4314)
- HBASE-26811 Secondary replica may be disabled for read incorrectly forever (#4310)
- HBASE-26673 Implement a shell command for change SFT implementation (#4113)
- HBASE-26838 Junit jar is not included in the hbase tar ball, causing … (#4223)
- HBASE-26586 Should not rely on the global config when setting SFT implementation for a table while upgrading (#4006)
- HBASE-26611 Changing SFT implementation on disabled table is dangerous (#4082)
- HBASE-26912 Bump checkstyle from 8.28 to 8.29 (#4293)
- HBASE-26871 addendum. use the jar command from JAVA_HOME (#4297)
- HBASE-26871 shaded mapreduce and shaded byo-hadoop client artifacts contain no classes (#4279)
- HBASE-26903 Bump httpclient from 4.5.3 to 4.5.13 (#4296)
- HBASE-26587 Introduce a new Admin API to change SFT implementation (#4030) (#4080)
- HBASE-26640 Reimplement master local region initialization to better work with SFT (#4111)
- HBASE-26690 Modify FSTableDescriptors to not rely on renaming when writing TableDescriptor (#4054)
- HBASE-26707: Reduce number of renames during bulkload (#4066) (#4122)
- HBASE-26483. [HBOSS] add support for createFile() and openFile(path) (#35)
- HBASE-27196: Enable code coverage reporting to SonarQube in hbase-filesystem (#36)
- HBASE-27076. [HBOSS] compile against hadoop 3.3.2+ only. (#34)
- HBASE-27042. Remove S3Guard awareness from HBoss
- HBASE-26786 [hboss] Limit synchronization from hot path of HBoss APIs (#32)
- HBASE-27135 [hbase-thirdparty] Bump checkstyle from 8.28 to 8.29 in /hbase-noop-htrace (#86)
- HBASE-27134 [hbase-thirdparty] Bump junit from 4.12 to 4.13.1 in /hbase-noop-htrace (#85)
- HBASE-27130 [hbase-thirdparty] Bump dependency versions (#83)
- HBASE-27133 Bump checkstyle from 8.28 to 8.29 in /hbase-unsafe (#84)
- HBASE-26893 [hbase-thirdparty] Upgrade jackson to 2.13.3 (#82)
- HBASE-26781 [hbase-thirdparty] Introduce the sun.misc.Signal delegati… (#79)
- HBASE-26773 [hbase-thirdparty] Introduce a hbase-unsafe module in hbase-thirdparty to remove the direct references of Unsafe in our main code base (#78)
- HBASE-26746 Update protobuf-java to 3.19.4 (#77)
- HBASE-26733 [hbase-thirdparty] Upgrade Netty to 4.1.73.Final (#75)
- HBASE-26732 [hbase-thirdparty] Update jackson (databind) to 2.13.1 (#74)
- HBASE-26592 Fix the broken shaded protobuf module (#70)
- HBASE-26506 Addendum add jersey-media-jaxb dependency (#69)
- HBASE-26504 Addendum upgrade extra-enforcer-rules to 1.4
- HBASE-26515 [hbase-thirdparty] Generate CHANGES.md and RELEASENOTES.md for 4.0.0 (#68)
- HBASE-26514 [hbase-thirdparty] Set version as 4.0.0 in prep for first RC (#67)
- HBASE-25868 [hbase-thirdparty] Shade jackson-jaxrs-json-provider for use with shaded jersey
- HBASE-25863 [hbase-thirdparty] Shade javax.ws.rs package for use with shaded Jersey
- HBASE-26503 [hbase-thirdparty] Bump guava version to 31.0.1-jre (#65)
- HBASE-26501 [hbase-thirdparty] Bump jetty version to 9.4.44.v20210927 (#61)

- HBASE-26506 [hbase-thirdparty] Bump jersey version to 2.35 (#64)
- HBASE-26505 [hbase-thirdparty] Bump commons-cli version to 1.5.0 (#63)
- HBASE-26504 [hbase-thirdparty] Bump maven plugin versions (#66)
- HBASE-26502 [hbase-thirdparty] Bump gson version to 2.8.9 (#60)
- HBASE-26499 [hbase-thirdparty] Bump netty version to 4.1.70.Final (#59)
- HBASE-25609 [hbase-thirdparty] Bump version to 4.0.0-SNAPSHOT on master branch (#57)
- HBASE-26500 [hbase-thirdparty] Bump protobuf version to 3.19.1 (#58)
- HBASE-26496 [hbase-thirdparty] Exclude the original protobuf-java jar when shading (#56)
- HBASE-25728 [hbase-thirdparty] Upgrade Netty library to >= 4.1.60 due to security vulnerability CVE-2021-21295 (#48)
- HBASE-26501 [hbase-thirdparty] Bump jetty version to 9.4.44.v20210927 (#61)

Fixed Issues in HDFS

Review the list of HDFS issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-27239: This commit reverts HDFS-13671, which was meant to reduce NameNode heap consumption, but it occasionally causes severe performance problems for some users, and that it does not go away without restarting the process. The revert will stabilize the NameNode performance for these users.

OPSAPS-63529: The users now can specify if they want the automatic rebootstrapping in case of HDFS incremental replications in the case of replication failure from the UI.

Fixed Issues in Apache Hive

Review the list of Hive issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-40730 PARQUET-1682: Forward compatibility for TIME/TIMESTAMP

PARQUET-1682 was backported to maintain forward compatibility for TIME/TIMESTAMP. This resolves the incompatibility between Hive's direct Parquet dependency and the transitive Parquet versions brought in by Iceberg.

CDPD-23454: Thrift version upgraded to 0.14.1 after this patch to avoid CVE.

This issue has been fixed.

CDPD-39708: This patch uses guava dependency version from cdpd repro.

This issue has been fixed.

CDPD-45199: Disable Strict Filtering and excluded reload4j jars from some places to prepare for hadoop to switch to reload4j.

This issue has been fixed.

CDPD-41660: Prevent creating embedded HMS instance at compaction if remote metastore uri is configured.

This issue has been fixed.

CDPD-31048: This patch provides a way to create only external tables in a database using a hive config.

This issue has been fixed.

OPSAPS-62325: set hive.optimize.dynamic.partition.hashjoin and hive.convert.join.bucket.mapjoin.tez to default true.

This issue has been fixed.

OPSAPS-63456: This fix allows CM to run Hive Metastore Upgrade tool for every upgrade. If the schema is already upto date, this becomes a no-op for HMS.

This issue has been fixed.

OPSAPS-65115: Added HiveUpgradeHandler for every release.

This issue has been fixed.

Apache Patch Information

- HIVE-22957
- HIVE-23482
- HIVE-24590
- HIVE-25303
- HIVE-25313
- HIVE-25632
- HIVE-25635
- HIVE-25724
- HIVE-25726
- HIVE-25800
- HIVE-25803
- HIVE-25826
- HIVE-25829
- HIVE-26055
- HIVE-26071
- HIVE-26109
- HIVE-26172
- HIVE-26230
- HIVE-26270
- HIVE-26322
- HIVE-26350
- HIVE-26394
- HIVE-26419
- HIVE-26488
- HIVE-26579
- HIVE-26613

Technical Service Bulletins

TSB 2023-627: IN/OR predicate on binary column returns wrong result

For the latest update on this issue, see the corresponding Knowledge article: TSB 2023-627: IN/OR predicate on binary column returns wrong result

TSB 2023-653: Cleaner causes data loss when processing an aborted dynamic partitioning transaction

For the latest update on this issue, see the corresponding Knowledge article: TSB 2023-653: Cleaner causes data loss when processing an aborted dynamic partitioning transaction

Fixed Issues in Hive Warehouse Connector

There are no fixed issues for HWC in Cloudera Runtime 7.2.16.

Fixed Issues in Hue

Review the list of Hue issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-41134: Hue Job Browser displays Impala query history in UTC timezone

Earlier, Hue used to display Impala query history in the UTC timezone, which caused confusion to users in other time zones. This issue has been fixed. Impala query history is now displayed as per the server timezone.

CDPD-40785: Export All feature in Hue needs a location to be passed but fails to export as the managed table does not accept location

Earlier, exporting query results using the "Export ALL" feature used to fail because the managed tables did not accept a location other than the one set in Hive. This issue has been fixed.

CDPD-41658: Unable to submit Oozie workflow from HDFS file browser

This issue has been fixed.

CDPD-25423: Downloaded query logs contain the error.json file, but do not fetch container logs

The debug bundles ZIP file no longer contains the logs. You can customize, generate, and view the logs by going to Cloudera Manager Diagnostics Logs .

CDPD-41666: Hue Importer does not work when the Impala editor is not available

This issue has been fixed.

CDPD-29285: Deselecting the Enable LDAP TLS option in Cloudera Manager does not work as expected

Earlier, when you deselected the Enable LDAP TLS option in Cloudera Manager Hue Configuration to enable LDAP authentication using unsecure LDAP (ldap:// instead of ldaps://), the authentication failed with the following error: Caught LDAPError while authenticating uldap: UNAVAILABLE({'info': '00000000: LdapErr: DSID-0C090F77, comment: Error initializing SSL/TLS, data 0, v23f0', 'desc': 'Server is unavailable'}. This issue has been fixed by forcing Hue to use the default value of the use_start_tls property (which is false) irrespective of the value present in ldap_cert property.

CDPD-44349: Oozie MapReduce action does not accept s3a and abfs paths as a file path in the Jar name field

This issue has been fixed.

CDPD-27884: Hue does not work with ID Broker HA, resulting in an error when you try to access S3/ADLS

This issue has been fixed.

CDPD-44832: Rerunning Oozie workflow fails

This issue has been fixed.

CDPD-44973: No attribute 'StringIO' error while loading snappy-compressed Hbase tables while the Avro module is not loaded

This issue has been fixed.

CDPD-41497: Unable to upload files to folders on S3/ABFS that contain non-ASCII characters in the folder name

This issue has been fixed.

Downloading Impala query results containing special characters in CSV format fails with ASCII codec error

This issue has been fixed.

Impala SELECT table query fails with UTF-8 codec error

This issue has been fixed.

Hue Load Balancer role fails to start after upgrade to Cloudera Runtime 7 or you get the "BalancerMember worker hostname too long" error

Hue supports creating Hue role hostnames of more than 64 characters. There is no longer a restriction of 64 characters for the “BalancerMember Route” property.

Fixed Issues in Apache Impala

Review the list of Impala issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-44771: Fixes hdfs_zone_alias_conf used with an Ozone storage path.

CDPD-43370: Improvements for Ozone: enables ofs as a default filesystem; fixes query performance when executors are co-located with Ozone datanodes.

CDPD-42618: Add STRING overloads for functions that consume time of day.

CDPD-41661: Fix nightly build issue targeting S3 filesystem.

CDPD-41346: Impala-shell bundled with CDP can now be used with Python 3.0-3.8.

CDPD-35771: Fix ORC Async IO crash due to unknown orc::StreamKind.

CDPD-34940: impala-shell can be used with Python 3 by installing the latest release from PyPI at https://pypi.org/project/impala-shell.

CDPD-31901: Fixes query performance when executors are co-located with Ozone datanodes.

CDPD-23458: Upgrade CPP thrift compiler version to thrift-0.16.0.

CDPD-22354: Added support for authentication via JWT for Impala.

OPSAPS-65040: Improved performance of ImpalaFileFormatAnalysisRule.

Technical Service Bulletin

TSB 2022-543: Impala query with predicate on analytic function may produce incorrect results

For the latest update on this issue see the corresponding Knowledge article: TSB 2022-543: Impala query with predicate on analytic function may produce incorrect results

Apache Patch Information

- IMPALA-11586
- IMPALA-11457

Fixed Issues in Apache Kafka

Review the list of Apache Kafka issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-40985: Expose log directory total and usable space through the Kafka API (KAFKA-13958 backport)

This is a backported improvement, see KIP-827 and KAFKA-13958 for more information.

CDPD-35680: Remove the verification of the existence of the JWT "sub" claim to extend compatibility with OAuth providers (KAFKA-13730 backport)

This is a backported improvement, see KAFKA-13730 for more information.

Topics created with the kafka-topics tool are only accessible by the user who created them when the deprecated --zookeeper option is used

The --zookeeper option has been removed from the kafka-topics tool. As a result, encountering this issue is no longer possible. Use the --bootstrap-server option instead.

Certain Kafka command line tools require direct access to ZooKeeper

There are no longer any Cloudera supported Kafka command line tools that require direct ZooKeeper access or require the usage of the --zookeeper option. Use the --bootstrap-server option instead.

CDPD-39422: Consumer polls for coordinator in tight loop (KAFKA-13917 backport)

The consumer uses a backoff period controlled by retry.backoff.ms to wait before polling for the coordinator if the previous poll was unsuccessful. This is a backported improvement, see KAFKA-13917 for more information.

CDPD-39354: Kafka Connect connectors and tasks fail to start

Kafka Connect initializes the Secrets Storage right at startup. As a result, timing issues no longer occur in the configuration resolution of connectors and tasks.

CDPD-45958: Kafka client JAAS override policy validation is incorrect

The JAAS override filter policy now correctly filters based on the specified rules and does not refuse JAAS configurations because of unknown fields.

CDPD-44252: Exception during normal operation in MirrorSourceTask causes the task to fail instead of shutting down gracefully

Stopping the read of offsets in a worker of a MirrorSourceTask will now cause a graceful shutdown and the task can be restarted automatically at a later point.

CDPD-39391: Amazon S3 Sink fails when validating bucket names

Due to an issue with the AWS S3 bucket name validation of the Amazon S3 Sink connector, the connector encountered an exception when validating bucket names. This issue is now fixed.

OPSAPS-64606: Authorization issues if Kafka Connect is not installed

If the Kafka Connect role is not present on the cluster, then a Ranger policy (connect internal - t opic) is created with default, non-empty topic names. As a result, the Ranger policy include list cannot be empty and will not have any side effects on other Kafka operations.

OPSAPS-63526: Kafka's Ranger related log4j2 config generation is not idempotent

If Kafka fails to start up due a configuration change, the original cause for the startup failure is no longer replaced with an error related to log4j2 configuration after the first auto-attempted restart.

OPSAPS-63640: Monitoring a high number of Kafka producers might cause Cloudera Manager to slow down and run out of memory

This issue is fixed.

Apache patch information

- KAFKA-14281

Fixed Issues in Apache Knox

Review the list of Knox issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-45191: KNOX-2810 Fixed an issue causing login failures when there were special characters in the password.

CDPD-43467: Upgrade aws-java-sdk to 1.12.261+ due to CVE-2022-31159.

CDPD-42850: Fixed an issue causing the remote alias service to regenerate pac4j password at each startup.

CDPD-41589: In previous versions, Knox's token impersonation feature was not working together with the HadoopAuth authentication provider. Now, this is fixed, Knox tokens can be generated on behalf of other users regardless of the authentication mechanism before the service.

CDPD-41495: Fixed user-auth-test command in KnoxCLI after upgrading to Shiro 1.6.

CDPD-41440: Fixed an issue where IDBroker would not distinguish between MSIs with different case. Going forward MSI names are case insensitive.

CDPD-40729: From now on, in Knox's HadoopGroupProvider, the gateway-level CENTRAL_GROUP_CONFIG_PREFIX prefixed parameters are added together with any custom provider-level parameters into the final HadoopGroupProvider identity assertion filter of the generated web application.

CDPD-40520: Upgraded mysql-connector-java to 8.0.29 due to CVE-2022-21363, CVE-2021-2471.

CDPD-40354: ATLAS, ATLAS-API, and SCHEMA-REGISTRY services do not need special replayBufferSize configuration for large data upload, Knox makes them work OOTB.

CDPD-37025: KNOX-2736 Added retry logic to knox client.

CDPD-36413: Knox gateway and idbroker startup time improvements were added.

CDPD-35951: Added support for SAML keystore type in Knox.

CDPD-24808: When multiple instances of Schema Registry are running, Knox will use round-robin to forward the requests.

OPSAPS-61474: Knox's data/applications folder gets recreated every time Knox starts.

OPSAPS-62573: CSD code for handling log4j2 config files for Knox.

OPSAPS-64309: Modified the Knox 7.1.8 CSD to include the WebHDFS work-around for FIPS-enabled clusters.

OPSAPS-64387: When multiple instances of Schema Registry are running, Knox will use round-robin to forward requests.

Apache patch information

- KNOX-2810
- KNOX-2733
- KNOX-2747
- KNOX-2770
- KNOX-2782
- KNOX-2757
- KNOX-2736

Technical Service Bulletins

TSB 2023-630: Apache Knox - Server-side Request Forgery in host parameter

For the latest update on this issue see the corresponding Knowledge article: TSB 2023-630: Apache Knox - Server-side Request Forgery in host parameter

Fixed Issues in Apache Kudu

Review the list of Apache Kudu issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-39418: Removed logredactor jar from plugins-common as ranger-kudu subprocess fails to initialize

Added logredactor dependency in ranger-rms-hive plugin so that it can package into ews/lib folder while building ranger-rms tar file.

CDPD-40027: Ranger plugin unable to start up

Removed logredactor jar from plugins-common as ranger-kudu subprocess fails to initialize. Added logredactor dependency in ranger-rms-hive plugin so that it can package into ews/lib folder while building ranger-rms tar file.

KUDU-1959: Fix the counter in StartupProgressStepsRemainingMetric()

Before this fix the counter in StartupProgressStepsRemainingMetric() was incremented twice if tablets are not processed during the startup of a tablet server.

KUDU-2218: SSL3_WRITE_PENDING TlsSocket error

KUDU-3306: String column types in range partitions lead to issues while copying tables

KUDU-3346: Fix rebalancer tool fails to run with '--ignored_tservers'

Prior to this patch the validity of 'ignored_tservers' was checked when 'BuildClusterinfo', which leads to a failure when the 'raw_info' only contains information of tservers on a specific location. This patch fixes it by moving the parameter validity check into 'KsckResultsToClusterRawInfo', because ksck results contain original cluster information.

KUDU-3384: DRS-level scan optimization leads to failed scans

KUDU-3401: Unable to query Kudu tables from Hive with Kudu HMS Integration enabled

KUDU-3404: glog 0.6.0 increases the TLS usage of libkudu_client.so substantially

Apache Patch Information

- KUDU-75
- KUDU-1260
- KUDU-1620
- KUDU-1885
- KUDU-1921
- KUDU-3297
- KUDU-3308
- KUDU-3344

Fixed issues in Livy

Review the list of Apache Livy issues that are resolved in Cloudera Runtime 7.2.16.
OPSAPS-63997:Added support for Livy HA and multiple instance detection.

Fixed Issues in Ozone

Review the list of Ozone issues that are resolved in Cloudera Runtime 7.2.16.
CDPD-30040: ofs input path parsing has been improved to avoid URISyntaxException.

CDPD-29894: Users' ofs trash folder path will now use that user's short user name instead (conversion rule according to hadoop.security.auth_to_local), rather than the full Kerberos principal which could differ when the same user logs in from different terminals on different nodes.

OPSAPS-57827: If any Ozone Manager or Storage Container Manager hasn't been finalized after upgrade, its canary indicator will turn yellow.

OPSAPS-64016: Ozone service status will now reflect Ozone Manager, Storage Container Manager and Ozone DataNode health status correctly.

Apache patch information

Apache patches in this release.

- HDDS-5502
- HDDS-5041

Fixed Issues in Apache Oozie

Review the list of Oozie issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-40326: mapreduce.job.acl-view-job property in Oozie workflow.xml not taking full effect.

This issue is resolved.

CDPD-29098: Oozie - Replace log4j 1.x with reload4j.

This issue is resolved.

CDPD-43343: Oozie log streaming bug when log timestamps are the same on multiple Oozie servers.

This issue is resolved.

CDPD-43192: Added additional HBase Jars to sharelib to support proper HBase interaction.

CDPD-41425: The following issue was fixed: OOZIE-3254 LAST_ONLY and NONE execution modes: possible OutOfMemoryError when there are too many coordinator actions to materialize.

CDPD-41328: Oozie incorrectly deleted the action dir base instead of the action dir belonging to the Job and thus causing intermittent failures. This has been fixed.

CDPD-39134: OOZIE-3661: Oozie cannot handle environment variables with key=value content.

CDPD-34593: Oozie's Spark action was unable to send lineage information to Atlas. We've added a new authentication / credential type called KafkaCredentials which will obtain a delegation token from Kafka. For more information please see the "Action Authentication" documentation in the official Oozie documentation which is accessible from the Oozie UI.

CDPD-30246: There were Jars missing from Oozie's Sqoop sharelib when a user wanted to import from RDBMS to HDFS/Hive into ORC file format or into HBase. The missing Jars are added now.

CDPD-27164: Oozie will no longer use the LoadBalancer to issue a callback notification, but instead it will try all available Oozie instances one-by-one. If the callback succeeded against one of the Oozie instances, then we will not try the other ones. This way the LoadBalancer will not be used for such purposes.

OPSAPS-63816: Cloudera Manager will provide the address of all Oozie server instances as a configuration to all Oozie instances. This will be then used by Oozie's callback mechanism so that instead of making the callback through the LoadBalancer in HA mode, the callback will be attempted through each Oozie instance, and if one of them succeeds, then we stop. This way we'll no longer use the LoadBalancer, and make the callback mechanism safer by not having a middle-man.

Apache patch information

- OOZIE-3666
- OOZIE-3254
- OOZIE-3535
- OOZIE-3661

Fixed Issues in Phoenix

Review the list of Phoenix issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-40194: The AccessDeniedException is fixed in Hue while creating Phoenix SYSTEM tables. Switched the impersonation method to the supported one for Phoenix driver.

CDPD-39117: If AND and OR filters exist in the query while converting to the Phoenix filters from Spark filters, no parentheses are added to the OR filters. This leads to a wrong filter and finally a wrong data. PHOENIX-6683 fixes this issue by surrounding the OR filters with parentheses while converting Spark filters to Phoenix expressions.

CDPD-35717: CALCITE-903 introduces a transparent reconnection feature, which opens a new server-side connection in case it is expired from the server-side connection cache.

Although this is convenient for most read-only analytical workloads, this can cause a number a problems, including data loss for transactional connections. This patch disables the transparent reconnect feature by default, and also adds the transparent_reconnection property. The feature is re-enabled when the property is set to true.

Apache Patch Information

- CALCITE-5009 Transparent JDBC connection re-creation may lead to data loss
- [CALCITE-4752] PreparedStatement#SetObject() fails for BigDecimal values
- [CALCITE-4676] Avatica client leaks TCP connections
- [CALCITE-3822] Source distribution must not contain fonts under SIL OFL 1.1 license (category B)
- PHOENIX-2704] Multilingual decoded problem (DonnyZone)
- PHOENIX-5894 Table versus Table Full Outer join on Salted tables not … (#1395)
- PHOENIX-6721 CSV bulkload tool fails with FileNotFoundException if --output points to the S3 location
- PHOENIX-5894 Table versus Table Full Outer join on Salted tables not working
- PHOENIX-6646 System tables are not upgraded after namespace migration
- PHOENIX-6611 Fix IndexTool -snap option and set VERIFIED in PhoenixIndexImportDirectReducer
- PHOENIX-6601 Fix IndexTools bugs with namespace mapping
- PHOENIX-6427 Create sequence fails in lowercase schema
- PHOENIX-6662 Failed to delete when PK with DESC with IN clause (#1509)
- PHOENIX-6659 RVC query fix for variable column + fixed column (#1507)
- PHOENIX-6773 PhoenixDatabaseMetadata.getColumns() always returns null COLUMN_DEF
- PHOENIX-6766 Fix failure of sqlline due to conflicting jline dependency pulled from Hadoop 3.3.
- PHOENIX-6758 During HBase 2 upgrade PHOENIX Self healing task fails to create server side connection before reading SYSTEM.TASK (#1478)
- PHOENIX-6753 Update default HBase 2.4 version to 2.4.13
- PHOENIX-6755 SystemCatalogRegionObserver extends BaseRegionObserver which doesn't exist in hbase-2.4 branch (#1472)
- PHOENIX-6725 : ConcurrentMutationException when adding column to table/view (addendum)
- PHOENIX-6725 : ConcurrentMutationException when adding column to table/view (#1452)
- PHOENIX-6734 Revert default HBase version to 2.4.10
- PHOENIX-5534 - Cursors With Request Metrics Enabled Throws Exception (#1451)
- PHOENIX-6710 Revert PHOENIX-3842 Turn on back default bloomFilter for Phoenix Tables (#1436)
- PHOENIX-6705 PagedRegionScanner#next throws NPE if pagedFilter is not initialized. (#1433)
- PHOENIX-6699 Phoenix metrics overwriting DefaultMetricsSystem in RegionServers (addendum: set up hbase prefix for ITs)
- PHOENIX-6699 Phoenix metrics overwriting DefaultMetricsSystem in RegionServers
- PHOENIX-6665 PreparedStatement#getMetaData() fails on parametrized "select next ? values for SEQ" (addendum: add test)
- PHOENIX-6665 PreparedStatement#getMetaData() fails on parametrized "select next ? values for SEQ"
- PHOENIX-6661 Sqlline does not work on PowerPC linux
- PHOENIX-6646 System tables are not upgraded after namespace migration
- PHOENIX-6579 ACL check doesn't honor the namespace mapping for mapped views.
- PHOENIX-6596 Schema extraction double quotes expressions, resulting in un-executabe create statements
- PHOENIX-5865 Column that has default value can not be correctly indexed
- PHOENIX-6615 The Tephra transaction processor cannot be loaded anymore.
- PHOENIX-6611 Fix IndexTool -snap option and set VERIFIED in PhoenixIndexImportDirectReducer
- PHOENIX-6618 Yetus docker image cannot be built as openjdk 11.0.11 is no longer available
- PHOENIX-6601 Fix IndexTools bugs with namespace mapping

- PHOENIX-6583 Inserting explicit Null into a (fixed length) binary field is stored as an array of zeroes
- PHOENIX-6528 Fix view index read repair for the pks with variable length
- PHOENIX-6507 DistinctAggregatingResultIterator should keep original tuple order of the AggregatingResultIterator
- PHOENIX-6498 Fix incorrect Correlated Exists Subquery rewrite when Subquery is aggregate
- PHOENIX-6578 sqlline.py cannot be started from source tree
- PHOENIX-6574 Executing "DROP TABLE" drops all sequences
- PHOENIX-6568 NullPointerException in PHOENIX-queryserver-client not in phoenix-client-hbase
- PHOENIX-6563 Unable to use 'UPPER'/'LOWER' together with 'IN'
- PHOENIX-5072 Cursor Query Loops Eternally with Local Index, Returns Fine Without It
- PHOENIX-6534 Upgrades from pre 4.10 versions are broken
- PHOENIX-6486 Phoenix uses inconsistent chronologies internally, breaking pre-Gregorian date handling
- PHOENIX-6506 : Tenant Connection is not able to access/validate Global Sequences (#1261)
- PHOENIX-6413 Having cannot resolve alias (#1168)
- PHOENIX-6476 Index tool when verifying from index to data doesn't correctly split page into tasks (#1240) (#1248)
- PHOENIX-6762 Phoenix QueryServer cannot run correctly with python 3.8+
- PHOENIX-6727 get_view_names() returning empty list
- PHOENIX-6661 Sqlline does not work on PowerPC linux
- PHOENIX-4646 Update to Jetty 9.4.42.v20210604
- PHOENIX-4196] Consume all data from client before replying with HTTP/401
- PHOENIX-6841 Depend on omid-codahale-metrics
- PHOENIX-6800 Remove superfluous semicolon for import statement in UncoveredLocalIndexRegionScanner (#1512)
- PHOENIX-6798 Eliminate unnecessary reversed scan for AggregatePlan (#1511)
- PHOENIX-6711 Add support of skipping the system tables existence check during connection initialisation and create new table result iterator which doesn't require fetch meta data of table
- PHOENIX-6480 Move phoenix-tool and add support with generating default properties (#1245)
- PHOENIX-6509 PHOENIX-4424 Allow users to create DEFAULT and HBASE Schema (Uppercase Schema Names) (#1263)
- PHOENIX-6451 Update joni and jcodings versions
- PHOENIX-3067 Phoenix metrics system should not be started in mini-cluster mode
- PHOENIX-6767 Traversing through all the guideposts to prepare parallel scans is not required for salted tables when the query is point lookup (#1493)
- PHOENIX-6751 Force using range scan vs skip scan when using large IN clause (#1496)
- PHOENIX-6771 Allow only "squash and merge" from GitHub UI
- PHOENIX-6653 Add upgrade tests based on HBase snapshots
- PHOENIX-6530 - changed the tenant id generation for uniform distribution load generator (#1453)
- PHOENIX-6708 Bump junit from 4.13 to 4.13.1
- PHOENIX-6697 log4j-reload4j is missing from phoenix-assembly
- PHOENIX-6690 Bump HBase 2.4 version to 2.4.11
- PHOENIX-6682 Jenkins tests are failing for Java 11.0.14.1
- PHOENIX-6686 Update Jackson to 2.12.6.1
- PHOENIX-6679 PHOENIX-6665 changed column name for CURRENT seqence values
- PHOENIX-6616 Alter table command can be used to set normalization_enabled=true on salted tables
- PHOENIX-6658 Replace HRegion.get() calls
- PHOENIX-6663 Use batching when joining data table rows with uncovered local index rows (#1403)
- PHOENIX-6501 Use batching when joining data table rows with uncovered global index rows (#1399)
- PHOENIX-6636 Replace bundled log4j libraries with reload4j (addendum: fix python scripts)
- PHOENIX-6656 Reindent NonAggregateRegionScannerFactory

- PHOENIX-6458 Using global indexes for queries with uncovered columns (#1256)
- PHOENIX-6636 Replace bundled log4j libraries with reload4j
- PHOENIX-6645 Remove unneccessary SCN related properties from SYSTEM tables on upgrade
- PHOENIX-6576 Do not use guava's Files.createTempDir()
- PHOENIX-6441 Remove TSOMockModule reference from OmidTransactionProvider (addendum:split TransactionServiceManager to avoid CNFE with -Dwithout.tephra)
- PHOENIX-6441 Remove TSOMockModule reference from OmidTransactionProvider
- PHOENIX-6638 Test suite fails with -Dwithout.tephra
- PHOENIX-6591 Update OWASP plugin to latest
- PHOENIX-6604 Allow using indexes for wildcard topN queries on salted tables
- PHOENIX-6582 Bump default HBase version to 2.3.7 and 2.4.8
- PHOENIX-6600 Replace deprecated getCall with updated getRpcCall (#1361) (#1356)
- PHOENIX-6594 Clean up vararg warnings flagged as errors by Eclipse
- PHOENIX-6592 PhoenixStatsCacheLoader uses non-deamon threads
- PHOENIX-6586 Set NORMALIZATION_ENABLED to false on salted tables
- PHOENIX-6561 : Allow pherf to intake phoenix Connection properties as argument. (#1322)
- PHOENIX-6577 phoenix_sandbox.py incompatible with python3
- PHOENIX-6472 In case of region inconsistency phoenix should stop gracefully
- PHOENIX-6344: CASCADE on ALTER should NOOP when there are no secondary indexes (#1135)
- PHOENIX-6555 Wait for permissions to sync in Permission tests
- PHOENIX-6548: Throw IOException instead of IllegalArgumentException when RS crashes during index rebuilds (#1329)
- PHOENIX-6557 Fix code problems flagged by SpotBugs as High priority
- PHOENIX-6556 Log INPUT_TABLE_CONDITIONS for MR jobs
- PHOENIX-6551 Bump HBase version to 2.4.6 and 2.2.7
- PHOENIX-6550 Upgrade jetty, jackson and commons-io
- PHOENIX-6546 BackwardCompatibilityIT#testSystemTaskCreationWithIndexAsyncRebuild is flakey
- PHOENIX-6526 Bump default HBase version on 2.3 profile to 2.3.6
- PHOENIX-6480 Move phoenix-tool and add support with generating default properties (#1245)
- PHOENIX-6519 Make SchemaTool work with lower case table and column names
- PHOENIX-6518 Implement SHOW CREATE TABLE SQL command
- PHOENIX-6454: Add feature to SchemaTool to get the DDL in specification (Addendum) (#1233)
- PHOENIX-6454: Add feature to SchemaTool to get the DDL in specification mode (#1229)
- PHOENIX-6515 Phoenix uses hbase-testing-util but does not list it as a dependency
- PHOENIX-6405 Disallow bulk loading into non-empty tables with global secondary indexes
- PHOENIX-6514 Exception should be thrown
- PHOENIX-6509 PHOENIX-4424 Allow users to create DEFAULT and HBASE Schema (Uppercase Schema Names) (#1263)
- PHOENIX-6770 set Log4j dependencies to provided
- PHOENIX-6683 Surround the OR filters with parentheses while convertin… (#73)
- PHOENIX-6694 Avoid unnecessary calls of fetching table meta data to region servers holding the system tables in batch oriented jobs in spark or hive otherwise those RS become hotspot
- PHOENIX-6450 Checkstyle creating warnings for line length > 80 but < 100

Fixed Issues in Parquet

Review the list of Parquet issues that are resolved in Cloudera Runtime 7.2.16.

Apache Patch Information

- PARQUET-2094

Fixed Issues in Apache Ranger

Review the list of Ranger issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-46781: Improve validation of condition expressions used in Ranger policies.

CDPD-46408: Upgrade spring security version to 5.7.5 as part of CVE fix.

CDPD-46309: Raz authorization for S3 and ALDS also will have audit metrics now. Issue here was wrong configuration prefix was sent and that resulted in not getting the right configuration values for authorization throughput metrics to be created.

CDPD-46256: Fixed Audit metrics not loading in new UI.

CDPD-46243: RAZ environment was failing to come up because of Audit metrics Initiation error. This was fixed on the RAZ.

CDPD-46233: Knox service was failing when Audit metrics was enabled. Fix was done to handle the CNF error in knox ranger plugin which took care of this error.

CDPD-45975: Audit metrics graphs were failing at the JPA level as different flavors of Database supported different predefined function. Fix is to use standard function which works across different databases.

CDPD-45874: Audit metric API issue is fixed by correcting the failing jpa query.

CDPD-45680: Replace log4j 1 with reload4j to fix the Log4j-1 EOL issue.

CDPD-45254:.env variables should be set before running docker-compose up.

CDPD-45116: Ranger admin user should able to change another user email after the upgrade.

CDPD-44810: updated resource signature of policy after modification of security zone name.

CDPD-44694: Change sync_source column datatype from varchar to text.

CDPD-44675: fixed url validation.

CDPD-44666: Policies that are maintained for the roles are not considered in the Ranger Policy reports and hence the report is not accurate. This fix is to generated the correct report.

CDPD-43941: Python client to test performance of CRUD operations on Ranger policy REST APIs.

CDPD-43873: Added validation for Validity Scheduler.

CDPD-43822: service creator user should able to create user of the ranger default policies.

CDPD-43792: The patch fixes the runtime complexity of ranger upgrade under heavy load(large no. of users and groups).

CDPD-43771: Improves the running time of java patch 55.

CDPD-43751: Improve java patch J10056 execution time while updating the large number of users.

CDPD-43465: Upgrade aws-java-sdk to 1.12.261 and azure-storage-blob to 12.18.0.

CDPD-43412: Fix ranger install script failure in python 3 env.

CDPD-42752: There is a change in external user 'status' (i.e x_portal_user tables column) which are getting synced into ranger admin, default ‘status’ value of synced users are getting set as 0(disabled) which was not the case in 7.1.4 This is the behaviour change between 7.1.4 and later versions.

Added change to mark external users status as enable(1). Written a java patch to update the status of existing external users.

CDPD-42607: Incremental Sync config parameter is read from config with the fix.

CDPD-41280: Fix Java patch J10033 and J10046 failure during ranger upgrade.

CDPD-41200: Show the alert only once if the resource lookup fails.

CDPD-41153: Updating the service config during upgrade which has unsupported access types e.g - others, solr_admin.

CDPD-40961: Opensearch Support on Ranger Admin.

CDPD-40268: 1. Fixed infinite loop in filtering out Ancestor resources 2. Filter out default-db from the mapped Hive resource if any other resource is also mapped. 3. Checking all mapped hive resources for access permissions 4. Added a config flag - ranger-rms.enable.database.sync (default true). If true, database level sync is enabled 5. Fixed ALTER_TABLE notification event processing 6. Added more debug messages to ChainedPlugin

CDPD-39931: Remove duplicate access types entries during the ranger policy creation.

CDPD-39803: Authmigrator Utility should not print irrelevant error messages.

CDPD-39594: Improved performance on the HBaseAuthorization request.

CDPD-39588: Exclude tag policies while transforming ranger policies through ranger policymigration module.

CDPD-39413: Print the skipped policy count while importing ranger policies through Sentry AuthMigrator tool.

CDPD-39412: Print the skipped policy count while importing ranger policies from the Sentry AuthMigrator tool.

CDPD-39360: Fix the serviceType ID and serviceType mapping of kafka in the ranger policymigration module.

CDPD-39359: Replace ElasticSearch to OpenSearch 1.3.2 in Ranger due to CVE. ElasticSearch cannot be upgraded due to Licensing issue hence it is replaced with OpenSearch 1.3.2.

CDPD-39594: Add python3 support in ranger install scripts.

CDPD-39319: Fix NullPointerException in get service REST call.

CDPD-39317: Updated atlas default audit filter to avoid auditing for atlas read-entity by nifi service user.

CDPD-39232: Hive table owner who create the tables full privilege.

CDPD-39208: Remove unused RDBMS tables used by Ranger admin service.

CDPD-39095: This JIRA is to verify Kafka Ranger plugin works correctly after upgrade of Kafka version to 3.1.

CDPD-35865: There is a change in external user 'status' (i.e x_portal_user tables column) which are getting synced into ranger admin, default ‘status’ value of synced users are getting set as 0(disabled) which was not the case in 7.1.4 This is the behaviour change between 7.1.4 and later versions.

Added change to mark external users status as enable(1). Written a java patch to update the status of existing external users.

CDPD-35628: If RangerRMS cannot renew it's ticket cache due to a KDC communication problem then it will not retry it and we'll see periodic "No ticket found in the cache" error messages. If that happens, then it won't have a valid Kerberos ticket it will not be able to communicate with other services, like HMS.

CDPD-35447: Added Ranger Support on Datahub HDFS.

CDPD-33606: Upgraded the Kylin version

CDPD-25938: As part of this change we have removed one policy item from storage policy with rangerlookup user which has unused access type.

CDPD-20527: Added API Documentation in the user profile drop-down. Click on that it opens a new window and loads swagger UI.

OPSAPS-62307: Ranger configurations now expose a safety-valve for authorization-migration-site.xml to allow users to configure required properties for custom configuration of properties which user can configure during migration of policies from Sentry to Ranger.

OPSAPS-62954: After the fix the default policies created in Ranger Admin should contain the actual configured service users and principal.

OPSAPS-63953: Removed default value JAVA_HOME variable and gave freedom to users to export the variable from their side. If not set users will see an error/notification to set the JAVA_HOME variable.

OPSAPS-64271: Ranger configurations now expose a safety-valve for authorization-migration-site.xml to allow users to configure required properties for custom configuration of properties which user can configure during migration of policies from Sentry to Ranger.

Technical Service Bulletins

TSB 2023-644: Microsoft Azure parent directory deletion

For the latest update on this issue, see the corresponding Knowledge Base article: TSB 2023-644: Microsoft Azure parent directory deletion.

Apache Patch Information

- CDPD-41188 : RANGER-3905
- CDPD-43934 : RANGER-3907
- Revert "RANGER-3840
- RANGER-3840
- Revert "RANGER-3768
- CDPD-41324: RANGER-3807
- RANGER-3790
- RANGER-3779
- RANGER-3768
- RANGER-3765
- RANGER-3763
- RANGER-3764
- CDPD-40486: Incorrect merging of RANGER-3548
- RANGER-3744
- RANGER-3736
- CDPD-39772: Backport RANGER-3717 RANGER-3299 RANGER-3716 RANGER-3732
- RANGER-3687
- RANGER-3211
- RANGER-3698
- RANGER-3389
- RANGER-2759
- RANGER-3784
- RANGER-3593
- RANGER-3780
- RANGER-3600

- RANGER-3752
- RANGER-3782
- RANGER-3735
- RANGER-3797
- RANGER-3793
- RANGER-3693
- RANGER-3795
- RANGER-3829
- RANGER-3861
- RANGER-3931
- RANGER-3857
- RANGER-3887
- RANGER-3888
- RANGER-3959
- RANGER-3960
- RANGER-3957
- RANGER-3956
- RANGER-3932
- RANGER-3914
- RANGER-3916
- RANGER-3912
- RANGER-3852
- RANGER-3886
- RANGER-3854
- RANGER-3735
- RANGER-3725

Fixed Issues in Schema Registry

Review the list of Schema Registry issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-20977: Add RAW Avro JSON Schema API for Hive Integration

Added new endpoints where the client can GET the actual schema text as a JSON document. These endpoints were added as a subresource (.../schemaText) to the existing schema version resource endpoints, corresponding to the "schemaText" property of those, for example, /api/v1/schemaregistry/schemas/{name}/versions/latest/schemaText.

CDPD-39885: SchemaRegistryResource.uploadFiles fails w/ Timeout waiting for connection from pool

When downloading a file from Schema Registry, the stream was not properly closed which occasionally caused issues. This has been fixed by always closing the stream.

CDPD-40758: Improve upon how SchemaRegistryClient is used in connectors

Fixed the issue when setting the "value.converter.serdes.protocol.version" in connectors caused the connector to fail on startup because the configuration property was not properly converted to a byte value.

CDPD-41592: Confluent import should handle input without an actual version

Lines will be ignored in Confluent import format where only the first part corresponding to a Cloudera SR "meta" is present and schema version data is missing (or only a "null" string is there).

CDPD-45920: KafkaJsonDeserializer does not support primitive type literals and array types

Fixed JSON deserializer's handling of non-object literal values (simple strings, numeric literals or arrays).

Fixed Issues in Cloudera Search

Review the list of Cloudera Search issues that are resolved in Cloudera Runtime 7.2.16.

Note: From Cloudera Runtime 7.2.16 and higher, new Cloudera Search fixed issues are listed under Apache Solr.

Fixed Issues in Apache Solr

There are no fixed issues for Solr in Cloudera Runtime 7.2.16.

Technical Service Bulletins

TSB-847: CVE-2025-30065 Apache Parquet vulnerability

On April 1, 2025, a critical vulnerability in the parquet-avro module of Apache Parquet (CVE-2025-30065, CVSS score 10.0) was announced.

Remediation for affected versions

The Cloudera Search release patched through the CDP updates for the public cloud and private cloud base.

Note: Cloudera will not provide remediation options for unsupported versions, and has not tested mitigations on unsupported versions. Customers are advised to upgrade to a supported product version. For more information, refer to the Support Lifecycle Policy page.

Vulnerability details

Exploiting this vulnerability is only possible by modifying the accepted schema used for translating Parquet files and subsequently submitting a specifically crafted malicious file.

Schema parsing in the parquet-avro module of Apache Parquet 1.15.0 and previous versions allows bad actors to execute arbitrary code. Attackers may be able to modify unexpected objects or data that was assumed to be safe from modification. Deserialized data or code could be modified without using the provided accessor functions, or unexpected functions could be invoked.

Deserialization vulnerabilities most commonly lead to undefined behavior, such as memory modification or remote code execution.

Action required - Mitigation for affected Cloudera products:

Until the upgrade with Apache Parquet 1.15.1 or higher is available:

1. Utilize a File Integrity Monitoring (FIM) solution. This allows administrators to monitor files at the filesystem level and receive alerts on any unexpected or suspicious activity in the schema configuration.
2. Monitor network activity for any transmission of Parquet files, and alert on any unexpected activity.
3. Be cautious with Parquet files from unknown or untrusted sources. If possible, do not process files with uncertain origin or that came from outside the organization.
4. Ensure that only authorized users have access to endpoints that ingest Parquet files.

For the latest update on this issue see the corresponding Knowledge Article: TSB 2025-847: Critical Apache Parquet vulnerability CVE-2025-30065

Fixed Issues in Spark

Review the list of Spark issues that are resolved in Cloudera Runtime 7.2.16.

CDPD-3783: Cannot create views (CREATE VIEW statement) from Spark.

CDPD-3293: Cannot create views (CREATE VIEW statement) from Spark.

CDPD-2650: Spark cannot write ZSTD and LZ4 compressed Parquet to dynamically partitioned tables

Fixed Issues in Apache Sqoop

Review the list of Sqoop issues that are resolved in Cloudera Runtime 7.2.16.
CDPD-43434: A safe-guard was put in place to make sure Sqoop always loads the correct logging related Jars independently from the classpath order.

CDPD-31522: Change log messages in Sqoop not to encourage users use --direct option.

This issue is resolved.

Apache patch information

None

Fixed Issues in Streams Messaging Manager

Review the list of Streams Messaging Manager issues that are resolved in Cloudera Runtime 7.2.16.
CDPD-33699: Remove "adjustTopicOverviewMetrics" from SMM

Removed the logic introduced in 7.1.4/7.2.2.0, where in case the topicMetrics (bytes in/bytes out/messages in) are smaller for a larger time period, the smaller timeperiod's metrics will be displayed. For example, if the metrics are smaller for 30 days and then for 7 days, the 7 day metrics would be used.

CDPD-35136: Data Explorer should allow long keys to be viewed

Fixed the issue with the keys being truncated in the Data Explorer in case they are too long on SMM UI. In case the message's key is too long to render, a "read more" button will appear next to the truncated key, initiating a popup which will show the full content of the key.

CDPD-35374: Incorrect key/value for large numbers

Fixed issue where large numbers were rounded in the Data Explorer in case LongDeserialiser is used for either the key or the value while viewing the content of a partition.

CDPD-35897: Remove partitionMetrics from /api/v1/admin/metrics/consumers/group/{groupId} response

Removed the "topicPartitionMetrics" field from the response of "/api/v1/admin/metrics/consumers/group/{groupId}" endpoint, since the underlying metric has been removed. Additionally, the API path version was bumped to v2.

CDPD-36420: List icons should not appear for the validation errors

Removed bullet point from validation errors on the connector creation form.

CDPD-39137: Connect topic tracking line is not showing on UI

On the Connectors page, topic lineage lines are not showing up, only the line starts and ends.

Now, Topic lineage lines are shown on the Connectors page.

CDPD-39778: Should not suggest flow.snapshot to be a secret

Streams Messaging Manager now does not allow the user to mark the flow.snapshot as sensitive data.

CDPD-39826: The Restart button for the ConnectorTasks is permanently disabled

Issue with the Restart button being permanently inactive on the ConnectorDetails page is fixed.

CDPD-39980: /api/v1/admin/auth/access throws NPE on unsecure environments

Fixed issue "/api/v1/admin/auth/access" throwing InternalServerError (Http code 500) when accessing the endpoint in a non-kerberized environment.

CDPD-40286: When using Oracle database, we can have an NPE during configuration

When using Oracle database, customer has the option to provide custom connection properties. Due to a bug, these connection properties were mandatory, which caused an error in case they were not provided.

CDPD-40758: Improve upon how SchemaRegistryClient is used in connectors

Fixed the issue when setting the "value.converter.serdes.protocol.version" in connectors caused the connector to fail on startup because the configuration property wasn't properly converted to a byte value.

CDPD-40871: SMM should fill "partitionMetrics" on "/api/v2/admin/metrics/aggregated/topics/{topicName}" even when partitionMetrics are empty

All the data that is available from the topic partitions are filled regardless whether metrics were fetched from Cloudera Manager or not.

CDPD-41069: Topic can be edited after selecting REPLICATION_STATUS as the alert attribute

Fixed the topic selection dropdown status in the alert editor after various UI events.

CDPD-41420: Schema version is not displayed in SMM when Avro value serializer is chosen

In the Data Explorer now it is possible to see all the schema versions associated with the given topic.

CDPD-41514: SMM displays CPU usage chart instead of CPU load

Fixed issue where SMM displays CPU usage chart instead of CPU load on the broker details page.

CDPD-41542: Consumer instance host field is empty

Fixed issue where consumer instance host field is empty in the SMM UI consumer details page.

CDPD-41553: Replication charts do not render on first visit

Fixed issue where replication charts do not render on first visit in the SMM topic details page.

CDPD-43387: Broker Details page does not show the Cloudera Manager button

The Cloudera Manager buttons that navigate to the broker resource within Cloudera Manager were not visible in previous releases. Now you can navigate from the SMM's broker view to the Cloudera Manager's broker view.

CDPD-43474: KConnect metrics cannot be scraped by prometheus in secure environments

Fixed Prometheus not being able to scrape Connect metrics from secure Kafka clusters. The "connect.prometheus.metrics.port" configuration was removed. For details, see *Kafka Connect property configuration in Cloudera Manager for Prometheus*.

CDPD-43962: Performance improvement on the Broker Details page

The '/api/v2/admin/metrics/aggregated/brokers/{brokerId}' endpoint, called every time while opening the broker details page on the UI, was excessively slow when a large number of topics and partitions were present. This is now fixed by fetching partition metrics in bulk.

Fixed Issues in Streams Replication Manager

Review the list of Streams Replication Manager issues that are resolved in Cloudera Runtime 7.2.16.

Rolling restart unavailable for SRM

Streams Replication Manager rolling restart support is added to Cloudera Manager. Streams Replication Manager can be restarted and upgraded without losing service availability.

Fixed Issues in Apache YARN and YARN Queue Manager

Review the list of YARN and YARN Queue Manager issues that are resolved in Cloudera Runtime 7.2.16.
COMPX-11380: Queue Manager displays an error stating that it is unable to complete a request.

COMPX-1451: Queue Manager does not support multiple ResourceManagers.

COMPX-1451: Queue Manager does not support multiple ResourceManagers.

COMPX-4992: Unable to switch to absolute mode after deleting a partition using YARN Queue Manage.

COMPX-5264: Unable to switch to Weight mode on creating a managed parent queue in Relative mode.

COMPX-5549: Queue Manager UI sets maximum-capacity to null when you switch mode with multiple partitions.

COMPX-5589: Unable to add new queue to leaf queue with partition capacity in Weight/Absolute mode.

COMPX-7586: Max Parallel Apps cannot be changed for root queue.

OPSAPS-50291: Environment variables HADOOP_HOME, PATH, LANG, and TZ are not on the allow list.

OPSAPS-52066: Stacks under Logs Directory for Hadoop daemons are not accessible from Knox Gateway.

OPSAPS-57067: Yarn Service in Cloudera Manager reports stale configuration yarn.cluster.scaling.recommendation.enable.

Apache patch information

- YARN-9997
- YARN-2710
- YARN-10850
- YARN-6221
- YARN-10869
- YARN-10727
- YARN-10790
- YARN-11126
- YARN-10910
- YARN-10915
- YARN-11023
- YARN-10997
- YARN-11024
- YARN-10907
- YARN-11303

Fixed Issues in Zeppelin

Review the list of Zeppelin issues that are resolved in Cloudera Runtime 7.2.16
OPSAPS-64691:[7.2.7->7.2.16] creation of 7.2.7 datahub is failing with : Initialize Zeppelin Notebook failed on Zeppelin Server.

Workaround is to manually add the ranger policy for zeppelin.

Apache patch information

- None

Fixed Issues in Apache ZooKeeper

Review the list of Zookeeper issues that are resolved in Cloudera Runtime 7.2.16.
OPSAPS-59080: On secure clusters, CM configures ZooKeeper to only allow JMX connections using TLS 1.2 encryption (on java 8) or TLS 1.2, 1.3 (on java 11)

Apache Patch Information

- ZOOKEEPER-3263: JAVA9/11 Warnings: Illegal reflective access in zookeeper's kerberosUtil (3.5)
- ZOOKEEPER-3652: Synchronize ClientCnxn outgoing queue flush on a stable internal value
- ZOOKEEPER-4477: Single Kerberos ticket renewal failure can prevent all future renewals since Java 9

Fixed Issues In Cloudera Runtime 7.2.16.1

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.1.

The following issue is resolved:

- HOTREQ-1264 - HotFix for CDPD-39592 CDPD-29225 for Public Cloud 7.2.14
- HOTREQ-1258 - Request for include OPSAPS-64187 and OPSAPS-65242 into 7.2.15.8

Fixed Issues In Cloudera Runtime 7.2.16.2

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.2.

The following issue is resolved:

- HOTREQ-1369 - Ranger S3 policy fails for read only or write only access.

Fixed Issues In Cloudera Runtime 7.2.16.3

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.3.

CDE

The following issue is resolved:

- TSB 2024-650: Arbitrary file deletion vulnerability in Apache Zeppelin

CDH

The following issue is resolved:

- HOTREQ-1347 Select query with LIMIT clause can fail if there are marker files like "_SUCCESS" and "_MANIFEST"
- HOTREQ-1287 Wrong results for partitioned Parquet table when files contain partition column
- HOTREQ-1330 add a way to reenable abfs readahead
- HOTREQ-1334 S3 Copy Optimization
- HOTREQ-1315 Upgrade node.js due to CVE-2022-35255, CVE-2022-43548 and CVE-2022-32212

- HOTREQ-1344 Hot fix JIRA CDPD-47077 for Public Cloud 7.2.15
- HOTREQ-1320 HOTFIX for Bug - Add delegation token support for long running spark job
- HOTREQ-1369 Ranger S3 policy fails for read only or write only access
- HOTREQ-1275 Hotfix for - IMPALA-11751
- HOTREQ-1244 HOTFIX Request for CDPD-46957
- HOTREQ-1274 HOTFIX REQ for issue HUE with KNOX unable to open workflow links in new tab with right click

CFM

- HOTREQ-1376 Fix flow downloading via knox for LGIM
- HOTREQ-1282 CFM - Parameter context inheritence fail during startup
- HOTREQ-1372 CaptureChangeMySQL processor fixes
- HOTREQ-1399 Ship NIFI-11363 to solve ENGESC-19490

CDPD-44232: Performance and security enhancements in Hue

Python 2 has reached the end of life and is no longer supported. Hue now uses Python 3 which makes use of critical bug fixes and Common Vulnerabilities and Exposures (CVE) fixes for many third-party software dependencies. The following changes have been made in the Hue codebase in this release of CDP Public Cloud:

- Python libraries such as django-auth-ldap, django-axes, djangorestframework-simplejwt, Mako, Markdown, python-ldap, django-babel, django-mako, django-cors-headers, djangorestframework, eventlet, sqlparse, and so on have been upgraded from Python 2.7 to Python 3.8.
- The Django server has been upgraded from version 1.11.29 to 3.2.15.
- Hue now uses Gunicorn as a front-end server. Previously, Hue used the CherryPy server.

These upgrades bring significant performance improvement and stability in query execution, uploading, and importing files to S3 or ABFS. Operating System, Python version, and Python module upgrades have resulted in a stable environment and fixed more than 800 security vulnerabilities.

Known issue: CDPD-54714

This is due to a missing configuration in Cloudera Manager. When Hue is enabled with Knox as authentication backend and Hue also in HA mode, all Hue instance's hostname should be added in knox_proxyhosts. This issue is known since Hue is still built with Python 2. This is not the issue related to recent Hue Python 3 build change.

Follow the procedure available in the Integrate Hue with Knox documentation.

Fixed Issues In Cloudera Runtime 7.2.16.200

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.200.

CDH

- HOTREQ-1274 HOTFIX REQ for issue HUE with KNOX unable to open workflow links in new tab with right click
- HOTREQ-1287 Wrong results for partitioned Parquet table when files contain partition column
- HOTREQ-1422 Need HIVE-26779 on Public Cloud runtime 7.2.15.10-1
- HOTREQ-1424 Accessing service via cdp-proxy-api failed with 404 with logs having SAXParseException
- HOTREQ-1420 distcp -update skips files of same size, name when transferring from Hdfs to S3
- HOTREQ-1433 Backport CDPD-55226 / HADOOP-18705 to 7.2.15
- HOTREQ-1443 HOTFIX for ENGESC-20387 (Yarn cluster overview pg refresh button issue)
- HOTREQ-1448 Hotfix request for ENGESC-20611

- HOTREQ-1444 HotFix for HIVE-27330
- TSB 2023-655: Apache Ranger (Ranger) S3 policies for READ or WRITE are not evaluated on RAZ-enabled CDP Public Cloud 7.2.16 environments

CFM

- HOTREQ-1397 Nifi CM metrics are not appear because of missing ranger policy
- HOTREQ-1396 Nifi ranger-nar not use hadoop downstream library
- HOTREQ-1374 Fix nifiregistry to use proper jdbc drivers
- HOTREQ-1423 ExecuteScript processor not supporting Module Directory for python
- HOTREQ-1427 Ship NIFI-10792 to 7.2.16 (CFM-2.2.6)

Known issue: CDPD-54714

This is due to a missing configuration in Cloudera Manager. When Hue is enabled with Knox as authentication backend and Hue also in HA mode, all Hue instance's hostname should be added in knox_proxyhosts. This issue is known since Hue is still built with Python 2. This is not the issue related to recent Hue Python 3 build change.

Follow the procedure available in the Integrate Hue with Knox documentation.

Fixed Issues In Cloudera Runtime 7.2.16.300

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.300.

CDH

- HOTREQ-1468 keytab does not exist error - ENGESC-20539
- HOTREQ-1414 Patches for Hive Iceberg for Customer LINE
- HOTREQ-1429 Request to backport CDPD-55677 into the next 7.2.15.x maintenance release
- HOTREQ-1461 HOTFIX for CDPD-48847

CFM

- HOTREQ-1459 SHIP NIFI-11653 Security fix for 7.2.17 (CFM-2.2.7), 7.2.16 (CFM-2.2.6) and 7.2.15 (CFM-2.2.5)
- HOTREQ-1458 Ship NIFI-11614 security fix for 7.2.16 (CFM-2.2.6) and 7.2.15 (CFM-2.2.5)
- HOTREQ-1396 Nifi ranger-nar not use hadoop downstream library
- HOTREQ-1481 Ship NIFI-11744 security fix for 7.2.17 (CFM-2.2.7), 7.2.16 (CFM-2.2.6) and 7.2.15 (CFM-2.2.5)
- HOTREQ-1488 SHIP NIFI-11334 to CFM-2.2.6.0 (7.2.16) and CFM-2.2.7.0 (7.2.17)

Known issue: CDPD-54714

This is due to a missing configuration in Cloudera Manager. When Hue is enabled with Knox as authentication backend and Hue also in HA mode, all Hue instance's hostname should be added in knox_proxyhosts. This issue is known since Hue is still built with Python 2. This is not the issue related to recent Hue Python 3 build change.

Follow the procedure available in the Integrate Hue with Knox documentation.

Fixed Issues In Cloudera Runtime 7.2.16.400

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.400.

CDH

- HOTREQ-1506 Schema Registry schema import must not deduplicate schemas
- HOTREQ-1515 Need HIVE-13288 on top of CDP 7.2.16
- HOTREQ-1440 Query failures with error "repeated binary array (STRING) is not a group"
- HOTREQ-1478 Backport CDPD-46799 and CDPD-57996 to 7.2.16.xxx
- HOTREQ-1485 GET_TABLES perf issue after upgrading from 7.2.12 to 7.2.16
- HOTREQ-1573 Hot fix JIRA CDPD-55511 for Public Cloud 7.2.16
- HOTREQ-1510 Hotfix request for IMPALA-11557 and IMPALA-11558
- HOTREQ-1408 [Hue][ABFS] List root directories in RAZ enabled env
- HOTREQ-1543 Backport CDPD-60778 to CDP 7.2.16 and 7.2.17
- TSB 2023-704: File corruption when downloading files larger than 1 MB from ABFS with Hue File Browser
- TSB 2023-703: Risk of Data Loss when using Hue S3 File Browser

CFM

- HOTREQ-1503 Ship CFM-3513 for 7.2.16 (CFM-2.2.6) and 7.2.17 (CFM-2.2.7)
- HOTREQ-1502 Ship CFM-3498 to 7.2.16 (CFM-2.2.6) and 7.2.17 (CFM-2.2.7)
- HOTREQ-1520 Ship NIFI-11854 for 7.2.16 (CFM-2.2.6) and 7.2.17 (CFM-2.2.7)
- HOTREQ-1540 Ship NIFI-11924 to 7.2.16 and 7.2.17
- HOTREQ-1539 Ship NIFI-11744 security fix for 7.2.16 and 7.2.17

Known issues

CDPD-61655: While using the Hue UI, users cannot create file or folder with non ascii char on Azure systems.

Avoid using Non-ASCII characters when creating a file or folder on Azure systems using the Hue UI.

Technical Service Bulletins

TSB 2023-667: HBase snapshot export failure can lead to data loss

For the latest update on this issue see the corresponding Knowledge article: TSB 2023-667: HBase snapshot export failure can lead to data loss.

Fixed Issues In Cloudera Runtime 7.2.16.500

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.500.

CDH

- HOTREQ-1590: Need hotfix for HIVE-21100 on CDP 7.2.15.0

CFM

- HOTREQ-1587: SHIP NIFI-12160 for 7.2.17, 7.2.16 and 7.2.15

Known issues

CDPD-61655: While using the Hue UI, users cannot create file or folder with non ascii char on Azure systems.

Avoid using Non-ASCII characters when creating a file or folder on Azure systems using the Hue UI.

Fixed Issues In Cloudera Runtime 7.2.16.600

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.600.

CDH

- HOTREQ-1571: Add configuration option to make container allocation prefer nodes without reserved containers.
- TSB 2024-723: Hue RAZ is using logger role to Read and Upload/Delete (write) files.

Known issues

CDPD-61655: While using the Hue UI, users cannot create file or folder with non ascii char on Azure systems.

Avoid using Non-ASCII characters when creating a file or folder on Azure systems using the Hue UI.

Fixed Issues in Cloudera Runtime 7.2.16.800

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.800.

CDPD-68335: Ranger Plugin support to use Solr ZooKeeper ClientConfig for writing audits to Solr when ZooKeeper SSL is enabled

Ranger plugin for writing audit information to Solr now supports ZooKeeper-Secure Sockets Layer (SSL) enabled connection.

CDPD-67600: Knox - Upgrade PostgreSQL version to 42.5.5/42.6.1/42.7.2 due to CVE-2024-1597

Upgraded the PostgreSQL version to 42.5.5/42.6.1/42.7.2 due to CVE-2024-1597.

CDPD-67060: Backport - ZEPPELIN-5929 - Spark Basic Feature Tutorial notebook data link is broken

The data link to the Spark Basic Features notebook was broken. This issue is now resolved.

CDPD-63760: Upgrade JGit version to 5.13.3.202401111512-r/6.6.1.202309021850/6.7.0.202309050840 due to CVE-2023-4759

Upgraded the JGit version to 5.13.3.202401111512-r/6.6.1.202309021850/6.7.0.202309050840 due to CVE-2023-4759

CDPD-63289: Upgrade amqp-client to 5.18.0 due to CVE-2023-46120

Upgraded the amqp-client version to 5.18.0 and above due to CVE-2023-46120.

CDPD-53844: HADOOP-18012. ABFS: Enable config controlled ETag check for Rename idempotency

Configuration controlled ETag check is now enabled for rename idempotency in Azure Blob File System (ABFS) driver .

Fixed Issues in Cloudera Runtime 7.2.16.900

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.16.900. This service pack was released on 30 Jul, 2024.

OPSAPS-70683: Cloudera Manager-7.9.0-h11 line agent unit test failures

Fixed line agent unit test failures in Cloudera Manager.

CDPD-72248: Zeppelin - Upgraded MomentJS to 2.29.4 due to CVE-2022-24785, CVE-2022-31129, and CVE-2017-18214

Upgraded MomentJS version to 2.29.4 due to CVE-2022-24785, CVE-2022-31129 and CVE-2017-18214.

CDPD-71580: Workaround needed for Bootbox due to CVE-2023-46998

Bootbox.js library was outdated. It is now removed and a new library Bootprompt is now used.

CDPD-71316: Backport HIVE-25501 to 7.2.16.x

A configurable filter is now provided to remove unused properties from PartitionDesc objects before MapWork serialization.

CDPD-70148: Backport Hadoop-18890

Removed the use of OkHttp in runtime code to simplify the dependencies when Hadoop does not use multiple third-party library to make HTTP calls.

CDPD-70004: IMPALA-12681 Some local file descriptors not released when using remote spilling

Fixed an issue where partially written temporary files were removed without releasing the file descriptors.

CDPD-68736: Upgraded OpenSearch to 1.3.15 due to CVE-2023-45807

Upgraded the OpenSearch version to 1.3.15 due to CVE-2023-45807.

CDPD-68692: Output from Hue shows NULL whereas Beeline works

Fixed an issue of misinterpreting certain returned values as NULL from Thrift API.

CDPD-68490: Zeppelin: Upgraded JLine to 3.25.1 due to CVE-2023-50572

Upgraded the JLine version to 3.25.1 due to CVE-2023-50572.

CDPD-68278: HWC - Upgrade Netty to 4.1.108.Final due to CVE-2024-29025

Upgraded the Netty version to 4.1.108 due to CVE-2024-29025.

CDPD-67599: Impala - Upgrade PostgreSQL to 42.5.5/42.6.1/42.7.2 due to CVE-2024-1597

Upgraded the PostgreSQL version to 42.5.5/42.6.1/42.7.2 due to CVE-2024-1597.

CDPD-67222, CDPD-67224: Ozone - Upgraded Spring Framework to 6.1.6/6.0.19/5.3.34 due to CVE-2024-22243, CVE-2024-22259, and CVE-2024-22262

Upgraded the Spring Framework version to 6.1.6/6.0.19/5.3.34 due to CVE-2024-22243, CVE-2024-22259, and CVE-2024-22262.

CDPD-64474: Data Catalog Profilers - Upgraded Logback to 1.2.13/1.3.14/1.4.14 due to CVE-2023-6378 and CVE-2023-6481

Upgraded the Logback version to 1.2.13/1.3.14/1.4.14 due to CVE-2023-6378 and CVE-2023-6481.

CDPD-64113: Upgraded Reactor-Netty to 1.0.39/1.1.13 due to CVE-2023-34062 and CVE-2023-34054

Upgraded the Reactor-Netty version to 1.0.39/1.1.13 due to CVE-2023-34062 and CVE-2023-34054.

CDPD-62164: Ranger backup should support different buckets

Ranger backup previously supported only one bucket. It now supports multiple buckets.

Known Issues In Cloudera Runtime 7.2.16

You must be aware of the known issues and limitations, the areas of impact, and workaround in Cloudera Runtime 7.2.16.

Known Issues in Apache Atlas

Learn about the known issues in Apache Atlas, the impact or changes to the functionality, and the workaround.

CDPD-48122: Support for the sortBy parameter under - api/atlas/admin/audits API is not available. Therefore, audit results under the administration audit page are working with limitations. You cannot sort the administration audit results.

None

CDPD-50239: The 'ATLAS_ENTITY_AUDIT_EVENTS' table increases its size considerably causing the disk space to shrink to a larger extent.

There are certain columns of the tables that are not configured to remove the old data that can cause disk to run out of space. For more information, see the section *Audit enhancements* in the mainline Atlas documentation for 7.2.16 version.

CDPD-46606: Performing Hive queries renders a notification for update data in the Hive table. Such a situation can lead to Kafka lag if there are many update query notifications.

None

CDPD-24089: Atlas creates HDFS path entities for GCP path and the qualified name of those entities does not have a cluster name appended.

None

CDPD-45642: When REST Notification server is down, messages from hooks are lost.

None

CDPD-46940: REST notification need to be disabled when running import scripts

None

CDPD-22082: ADLS Gen2 metadata extraction: If the queue is not cleared before performing Incremental extraction, messages are lost.

After successfully running Bulk extraction, you must clear the queue before running Incremental extraction.

CDPD-19996: Atlas AWS S3 metadata extractor fails when High Availability is configured for IDBroker.

If you have HA configured for IDBroker, make sure your cluster has only one IDBroker address in core-site.xml. If your cluster has two IDBroker addresses in core-site.xml, remove one of them, and the extractor must be able to retrieve the token from IDBroker.

CDPD-19798: Atlas /v2/search/basic API does not retrieve results when the search text mentioned in the entity filter criteria (like searching by Database or table name) has special characters like + - & | ! ( ) { } [ ] ^ " ~ * ? :

You can invoke the API and mention the search string (with special characters) in the query attribute in the search parameters.

ATLAS-3921: Currently there is no migration path from AWS S3 version 1 to AWS S3 version 2.

None

CDPD-12668: Navigator Spark lineage can fail to render in Atlas

As part of content conversion from Navigator to Atlas, the conversion of some spark applications created a cyclic lineage reference in Atlas, which the Atlas UI fails to render. The cases occur when a Spark application uses data from a table and updates the same table.

None

CDPD-11941: Table creation events missed when multiple tables are created in the same Hive command

When multiple Hive tables are created in the same database in a single command, the Atlas audit log for the database may not capture all the table creation events. When there is a delay between creation commands, audits are created as expected.

None

CDPD-11940: Database audit record misses table delete

When a hive_table entity is created, the Atlas audit list for the parent database includes an update audit. However, at this time, the database does not show an audit when the table is deleted.

None

CDPD-11790: Simultaneous events on the Kafka topic queue can produce duplicate Atlas entities

In normal operation, Atlas receives metadata to create entities from multiple services on the same or separate Kafka topics. In some instances, such as for Spark jobs, metadata to create a table entity in Atlas is triggered from two separate messages: one for the Spark operation and a second for the table metadata from HMS. If the process metadata arrives before the table metadata, Atlas creates a temporary entity for any tables that are not already in Atlas and reconciles the temporary entity with the HMS metadata when the table metadata arrives.

However, in some cases such as when Spark SQL queries with the write.saveAsTable function, Atlas does not reconcile the temporary and final table metadata, resulting in two entities with the same qualified name and no lineage linking the table to the process entity.

This issue is not seen for other lineage queries from spark:

```
create table default.xx3 as select * from default.xx2
insert into yy2 select * from yy
insert overwrite table ww2 select * from ww1
```

Another case where this behavior may occur is when many REST API requests are sent at the same time.

None

CDPD-11692: Navigator table creation time not converted to Atlas

In converting content from Navigator to Atlas, the create time for Hive tables is not moved to Atlas.

None

CDPD-11338: Cluster names with upper case letters may appear in lower case in some process names

Atlas records the cluster name as lower case in qualifiedNames for some process names. The result is that the cluster name may appear in lower case for some processes (insert overwrite table) while it appears in upper case for other queries (ctas) performed on the same cluster.

None

CDPD-10576: Deleted Business Metadata attributes appear in Search Suggestions

Atlas search suggestions continue to show Business Metadata attributes even if the attributes have been deleted.

None

CDPD-10574: Suggestion order doesn't match search weights

At this time, the order of search suggestions does not honor the search weight for attributes.

None

CDPD-9095: Duplicate audits for renaming Hive tables

Renaming a Hive table results in duplicate ENTITY_UPDATE events in the corresponding Atlas entity audits, both for the table and for its columns.

None

CDPD-7982: HBase bridge stops at HBase table with deleted column family

Bridge importing metadata from HBase fails when it encounters an HBase table for which a column family was previously dropped. The error indicates:

```
Metadata service API org.apache.atlas.AtlasClientV2$API_V2@58112
bc4 failed with status 404 (Not Found) Response Body
({""errorCode"":""ATLAS-404-00-007"",""errorMessage"":""Invalid
 instance creation/updation parameters passed :
hbase_column_family.table: mandatory attribute value missing in
 type hbase_column_family""})
```

None

CDPD-7781: TLS certificates not validated on Firefox

Atlas is not checking for valid TLS certificates when the UI is opened in FireFox browsers.

None

CDPD-6675: Irregular qualifiedName format for Azure storage

The qualifiedName for hdfs_path entities created from Azure blog locations (ABFS) doesn't have the clusterName appended to it as do hdfs_path entities in other location types.

None

CDPD-5933 and CDPD-5931: Unexpected Search Results When Using Regular Expressions in Basic Searches on Classifications

When you include a regular expression or wildcard in the search criteria for a classification in the Basic Search, the results may differ unexpectedly from when full classification names are included. For example, the Exclude sub-classifications option is respected when using a full classification name as the search criteria; when using part of the classification name and the wildcard (*) with Exclude sub-classifications turned off, entities marked with sub-classifications are not included in the results. Other instances of unexpected results include case-sensitivity.

None

CDPD-4762: Spark metadata order may affect lineage

Atlas may record unexpected lineage relationships when metadata collection from the Spark Atlas Connector occurs out of sequence from metadata collection from HMS. For example, if an ALTER TABLE operation in Spark changing a table name and is reported to Atlas before HMS has processed the change, Atlas may not show the correct lineage relationships to the altered table.

None

CDPD-4545: Searches for Qualified Names with "@" doesn't fetch the correct results

When searching Atlas qualifiedName values that include an "at" character (@), Atlas does not return the expected results or generate appropriate search suggestions.

Consider leaving out the portion of the search string that includes the @ sign, using the wildcard character * instead.

CDPD-3208: Table alias values are not found in search

When table names are changed, Atlas keeps the old name of the table in a list of aliases. These values are not included in the search index in this release, so after a table name is changed, searching on the old table name will not return the entity for the table.

None

CDPD-3160: Hive lineage missing for INSERT OVERWRITE queries

Lineage is not generated for Hive INSERT OVERWRITE queries on partitioned tables. Lineage is generated as expected for CTAS queries from partitioned tables.

None

CDPD-3125: Logging out of Atlas does not manage the external authentication

At this time, Atlas does not communicate a log-out event with the external authentication management, Apache Knox. When you log out of Atlas, you can still open the instance of Atlas from the same web browser without re-authentication.

To prevent access to Atlas after logging out, close all browser windows and exit the browser.

CDPD-1892: Ranking of top results in free-text search not intuitive

The Free-text search feature ranks results based on which attributes match the search criteria. The attribute ranking is evolving and therefore the choice of top results may not be intuitive in this release.

If you don't find what you need in the top 5 results, use the full results or refine the search.

CDPD-1884: Free text search in Atlas is case sensitive

The free text search bar in the top of the screen allows you to search across entity types and through all text attributes for all entities. The search shows the top 5 results that match the search terms at any place in the text (*term* logic). It also shows suggestions that match the search terms that begin with the term (term* logic). However, in this release, the search results are case-sensitive.

If you don't see the results you expect, repeat the search changing the case of the search terms.

CDPD-1823: Queries with ? wildcard return unexpected results

DSL queries in Advanced Search return incorrect results when the query text includes a question mark (?) wildcard character. This problem occurs in environments where trusted proxy for Knox is enabled, which is always the case for CDP.

None

CDPD-1664: Guest users are redirected incorrectly

Authenticated users logging in to Atlas are redirected to the CDP Knox-based login page. However, if a guest user (without Atlas privileges) attempts to log in to Atlas, the user is redirected instead to the Atlas login page.

To avoid this problem, open the Atlas Dashboard in a private or incognito browser window.

CDPD-922: IsUnique relationship attribute not honored

The Atlas model includes the ability to ensure that an attribute can be set to a specific value in only one relationship entity across the cluster metadata. For example, if you wanted to add metadata tags to relationships that you wanted to make sure were unique in the system, you could design the relationship attribute with the property "IsUnique" equal true. However, in this release, the IsUnique attribute is not enforced.

None

CDPD-76789: Creating tag with name description throws java.lang.ClassCastException

Creating classification with reserved names such as "name", "description", "owner", "version", "serviceType" and "options" can lead to HTTP 500 error in Apache Atlas.

Avoid creating classification with reserved names such as "name", "description", "owner", "version", "serviceType" and "options".

Known Issues in Apache Avro

Learn about the known issues in HDFS, the impact or changes to the functionality, and the workaround.

CDPD-23451: Remove/replace jackson-mapper-asl dependency.

Avro library depends on the already EOL jackson-mapper-asl 1.9.13-cloudera.1 that also contains a couple of CVEs. The jackson library is part of the Avro API so cannot be changed without a complete rebase.

None.

Known issues in Cruise Control

Learn about the known issues in HDFS, the impact or changes to the functionality, and the workaround.

Rebalancing with Cruise Control does not work due to the metric reporter failing to report the CPU usage metric

On the Kafka broker, the Cruise control metric reporter plugin may fail to report the CPU usage metric.

If the CPU usage metric is not reported, the numValidWindows in Cruise Control will be 0 and proposal generation as well as partition rebalancing will not work. If this issue is present, the following message will be included in the Kafka logs:

```
WARN com.linkedin.kafka.cruisecontrol.metricsreporter.CruiseCont
rolMetricsReporter:
          [CruiseControlMetricsReporterRunner]: Failed reporting
 CPU util.
```

```
java.io.IOException: Java Virtual Machine recent CPU usage is not
        available.
```

This issue is only known to affect Kafka broker hosts that have the following specifications:

- CPU: Intel(R) Xeon(R) CPU E5-2699 v4 @ 2.20GHz
- OS: Linux 4.18.5-1.el7.elrepo.x86_64 #1 SMP Fri Aug 24 11:35:05 EDT 2018 x86_64
- Java version: 8-18

Move the broker to a different machine where the CPU is different. This can be done by performing a manual repair on the affected nodes. For more information, see the Data Hub documentation.

Note: Cluster nodes affected by this issue are not displayed as unhealthy.

CDPD-47616: Unable to initiate rebalance, number of valid windows (NumValidWindows) is zero

If a Cruise Control rebalance is initiated with the rebalance_disk parameter and Cruise Control is configured to fetch metrics from Cloudera Manager (Metric Reporter is set to CM metrics reporter), Cruise Control stops collecting metrics from the partitions that are moved. This is because Cloudera Manager does not collect metrics from moved partitions due to an issue in Kafka (KAFKA-10320).

If the metrics are not available, the partition is considered invalid by Cruise Control. This results in Cruise Control blocking rebalance operations and proposal generation.

Configure Cruise Control to use to use the Cruise Control metrics reporter (default). This issue is not present if this metric reporter is used.

1. In Cloudera Manager, select the Cruise Control service.
2. Go to Configuration.
3. Find the Metric Reporter property.
4. Select the Cruise Control metrics reporter option.
5. Restart the Cruise Control service.

OPSAPS-68148: Cruise Control rack aware goal upgrade handler

The goal sets in Cruise Control, which include the default, supported, hard, self-healing and anomaly detection goals, might be overridden to their default value after a cluster upgrade if the goals have been customized.

Create a copy from the values of the goal lists before upgrading your cluster, and add the copied values to the goal lists after upgrading the cluster. Furthermore, you must rename any mentioning of com.linkedin.kafka.cruisecontrol.analyzer.goals.RackAwareGoal to com.linkedin.kafka.cruisecontrol.analyzer.goals.RackAwareDistributionGoal as Cruise Control will not be able to start otherwise.

Known Issues in Data Analytics Studio

Learn about the known issues in Data Analytics Studio, the impact or changes to the functionality, and the workaround.

- CDPD-49281: DAS WebApp logs are not captured in the var/logs/das/ directory, as expected.

  Workaround: To obtain the DAS WebApp logs, check the stderr.log file in the runtime process directory for the DAS WebApp.
- You may not be able to add or delete columns or change the table schema after creating a new table using the upload table feature.
- For clusters secured using Knox, you see the HTTP 401: Forbidden error message when you click the DAS quick link from Cloudera Manager and are unable to log into DAS.

  Workaround: The admin user will need to provide the DAS URL from the Knox proxy topology to the users needing access to DAS.
- The download logs feature may not return the YARN application logs on a Kerberized cluster. When you download the logs, the logs contain an error-reports.json file which states that no valid Kerberos tokens are available.

  Workaround: An admin user with access to the machine can use the kinit command as a hive user with hive service user keytabs and trigger the download.
- The task logs for a particular task may not be available in the task swimlane. And the zip file generated by download logs artifact may not have task logs, but instead contain an error-reports.json file with the error log of the download failures.
- You may not see any data for a report for any new queries that you run. This can happen especially for the last one day's report.

  Workaround:

  1. Shut down the DAS Event Processor.
  2. Run the following command from the Postgres server:

     ```
     update das.report_scheduler_run_audit set status = 'FAILED' where status
      = 'READING';
     ```
  3. Start the DAS Event Processor.
- On clusters secured with Knox proxy only: You might not be able to save the changes to the JDBC URL in the DAS UI to change the server interface (HS2 or LLAP) on which you are running your queries.
- You may be unable to upload tables or get an error while browsing files to upload tables in DAS on a cluster secured using Knox proxy.
- DAS does not parse semicolons (;) and double hyphens (--) in strings and comments.

  For example, if you have a semicolon in query such as the following, the query might fail: select * from properties where prop_value = "name1;name2";

  If a semicolon is present in a comment, then run the query after removing the semicolon from the comment, or removing the comment altogether. For example:

  ```
  select * from test; -- select * from test;
  select * from test; /* comment; comment */
  ```

  Queries with double hyphens (--) might also fail. For example:

  ```
  select * from test where option = '--name';
  ```
- You might face UI issues on Google Chrome while using faceted search. We recommend you to use the latest version of Google Chrome (version 71.x or higher).
- Visual Explain for the same query shows different graphs on the **Compose** page and the **Query Details** page.
- While running some queries, if you restart HSI, the query execution is stopped. However, DAS does not reflect this change and the queries appear to be in the same state forever.
- After a fresh installation, when there is no data and you try to access the Reports tab, DAS displays an "HTTP 404 Not Found" error.
- Join count does not get updated for tables with partitioned columns.

Known Issues in Apache HBase

Learn about the known issues in HDFS, the impact or changes to the functionality, and the workaround.

OpDB Data Hub cluster fails to initialize if you are reusing a cloud storage location that was used by an older OpDB Data Hub cluster

Workaround: Stop HBase using Cloudera Manager before deleting an Operational Database Data Hub cluster.

IntegrationTestReplication fails if replication does not finish before the verify phase begins

During IntegrationTestReplication, if the verify phase starts before the replication phase finishes, the test fails because the target cluster does not contain all of the data. If the HBase services in the target cluster does not have enough memory, long garbage-collection pauses might occur.

Workaround: Use the -t flag to set the timeout value before starting verification.

HDFS encryption with HBase

Cloudera has tested the performance impact of using HDFS encryption with HBase. The overall overhead of HDFS encryption on HBase performance is in the range of 3 to 4% for both read and update workloads. Scan performance has not been thoroughly tested.

Workaround: N/A

AccessController postOperation problems in asynchronous operations

When security and Access Control are enabled, the following problems occur:

- If a Delete Table fails for a reason other than missing permissions, the access rights are removed but the table may still exist and may be used again.
- If hbaseAdmin.modifyTable() is used to delete column families, the rights are not removed from the Access Control List (ACL) table. The portOperation is implemented only for postDeleteCo lumn().
- If Create Table fails, full rights for that table persist for the user who attempted to create it. If another user later succeeds in creating the table, the user who made the failed attempt still has the full rights.

Workaround: N/A

Apache Issue: HBASE-6992

Bulk load is not supported when the source is the local HDFS

The bulk load feature (the completebulkload command) is not supported when the source is the local HDFS and the target is an object store, such as S3/ABFS.

Workaround: Use distcp to move the HFiles from HDFS to S3 and then run bulk load from S3 to S3.

Apache Issue: N/A

CDPD-77399: HBase fails to register the servlet metrics and throws ClassNotFoundException: org.apache.hadoop.metrics.MetricsServlet

The MetricsServlet class is a Hadoop 2-based metric servlet unavailable in Hadoop 3 deployments.

Workaround: Ignore this WARN log message during HBase Master and RegionServer startup.

Apache Issue: HBASE-28315

Known Issues in HDFS

Learn about the known issues in HDFS, the impact or changes to the functionality, and the workaround.

OPSAPS-55788: WebHDFS is always enabled. The Enable WebHDFS checkbox does not take effect.

None.

Unsupported Features

The following HDFS features are currently not supported in Cloudera Data Platform:

- ACLs for the NFS gateway (HADOOP-11004)
- Aliyun Cloud Connector (HADOOP-12756)
- Allow HDFS block replicas to be provided by an external storage system (HDFS-9806)
- Consistent standby Serving reads (HDFS-12943)
- Cost-Based RPC FairCallQueue (HDFS-14403)
- HDFS Router Based Federation (HDFS-10467)
- More than two NameNodes (HDFS-6440)
- NameNode Federation (HDFS-1052)
- NameNode Port-based Selective Encryption (HDFS-13541)
- Non-Volatile Storage Class Memory (SCM) in HDFS Cache Directives (HDFS-13762)
- OpenStack Swift (HADOOP-8545)
- SFTP FileSystem (HADOOP-5732)
- Storage policy satisfier (HDFS-10285)

Technical Service Bulletins

TSB 2023-666: Out of order HDFS snapshot deletion may delete renamed/moved files, which may result in data loss

Cloudera has discovered a bug in the Apache Hadoop Distributed File System (HDFS) snapshot implementation. Deleting an HDFS snapshot may incorrectly remove files in the .Trash directories or remove renamed files from the current file system state. This is an unexpected behavior because deleting an HDFS snapshot should only delete the files stored in the specified snapshot, but not data in the current state.

In the particular HDFS installation in which the bug was discovered, deleting one of the snapshots caused certain files to be moved to trash and deletion of some of the files in a .Trash directory. Although it is clear that the conditions of the bug are (1) out-of-order snapshot deletion and (2) files moved to trash or other directories, we were unable to replicate the bug in other HDFS installations after executing similar test operations with a variety of different sequences. We also did not observe any actual data loss in our tests. However, there is a remote possibility that this bug may lead to data loss.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: TSB 2023-666: Out of order HDFS snapshot deletion may delete renamed/moved files, which may result in data loss

Known Issues in Apache Hive

Learn about the known issues in Hive, the impact or changes to the functionality, and the workaround.

TSB-732 2024: Incorrect results are generated by Hive JOIN when bloom filter is activated

The bloom filter implemented in HIVE-23880 was designed to enhance performance for queries with JOIN statements, where one small table and another significantly larger table is joined on partition keys. However, the bloom filter introduced an issue in Apache Hive (Hive), when dynamic semijoin redaction is involved that generates incorrect query results. This issue is corrected in HIVE-26655.

Upstream JIRA

Hive-23880(cause)HIVE-26655(fix)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: TSB 2024-732: Incorrect results are generated by Hive JOIN when bloom filter is activated

CDPD-40730: Parquet change can cause incompatibility

Parquet files written by the parquet-mr library in CDP 7.2.16, where the schema contains a timestamp with no UTC conversion will not be compatible with older versions of Parquet readers. The effect is that the older versions will still consider these timestamps as they would require UTC conversions and will thus end up with a wrong result. You can encounter this problem only when you write Parquet-based tables using Hive, and tables have the non-default configuration hive.parquet.write.int64.timestamp=true.

CDPD-60770: Beeline Authentication Issue with Special Characters in Passwords

When LDAP is enabled, users cannot authenticate with Beeline if the password contains a special character. For example, the following string fails:

```
beeline -u jdbc:hive2://<host>:<port>/<dbName>;user=user@XXX;password='R3G#xpXyoy1MOJb1'
```

Use the -p parameter to execute the Beeline command:

```
beeline -u jdbc:hive2://<host>:<port>/<dbName>; -n user@XXX -p 'R3G#xpXyoy1MOJb1'
```

CDPD-15518: ACID tables you write using the Hive Warehouse Connector cannot be read from an Impala virtual warehouse.

Read the tables from a Hive virtual warehouse or using Impala queries in Data Hub.

CDPD-10848: HiveServer Web UI displays incorrect data

If you enabled auto-TLS for TLS encryption, the HiveServer2 Web UI does not display the correct data in the following tables: Active Sessions, Open Queries, Last Max n Closed Queries

Known Issues in Hue

Learn about the known issues in Hue, the impact or changes to the functionality, and the workaround.

Known issues in 7.2.16

CDPD-54714: SSO does not work while logging in from the Hue UI

Due to a missing configuration in Cloudera Manager, SSO does not work when you have enabled Knox as an authentication backend and when Hue is in HA mode.

See Authenticating Hue users with Knox SSO.

CDPD-41136: Importing files from the local workstation is disabled by default

Cloudera has disabled the functionality to import files from your local workstation into Hue because it may cause errors. You may not see the Local File option in the Type drop-down menu on the Importer page by default.

You can enable the functionality to import files from your local workstation by specifying the following parameter in the Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini field using Cloudera Manager:

```
[indexer]
enable_direct_upload=true
```

CDPD-42619: Unable to import a large CSV file from the local workstation

You may see an error message while importing a CSV file into Hue from your workstation, stating that you cannot import files of size more than 200 KB.

Upload the file to S3 or ABFS and then import it into Hue using the Importer.

CDPD-43293: Unable to import Impala table using Importer

Creating Impala tables using the Hue Importer may fail.

If you have both Hive and Impala services installed on your cluster, then you can import the table using by selecting the Hive dialect from Tables Sources . If only Impala service is installed on your cluster, then go to Cloudera Manager Clusters Hue Configurations and add the following line in the Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini field:

```
[beeswax]
max_number_of_sessions=1
```

Technical Service Bulletins

TSB 2024-723: Hue RAZ is using logger role to Read and Upload/Delete (write) files

When using Cloudera Data Hub for Public Cloud (Data Hub) on Amazon Web Services (AWS), users can use the Hue File Browser feature to access the filesystem, and if permitted, read and write directly to the related S3 buckets. As AWS does not provide fine-grained access control, Cloudera Data Platform administrators can use the Ranger Authorization Service (RAZ) capability to take the S3 filesystem, and overlay it with user and group specific permissions, making it easier to allow certain users to have limited permissions, without having to grant those users permissions to the entire S3 bucket.

This bulletin describes an issue when using RAZ with Data Hub, and attempting to use fine-grained access control to allow certain users write permissions.

Through RAZ, an administrator may, for a particular user, specify permissions more limited than what AWS provides for an S3 bucket, allowing the user to have read/write (or other similar fine grained access) permissions on only a subset of the files and directories within that bucket. However, under specific conditions, it is possible for such user to be able to read and write to the entire S3 bucket through Hue, due to Hue using the logger role (which will have full read/write to the S3 bucket) when using Data Hub with a RAZ enabled cluster. This problem also can affect the Hue service itself, by affecting proper access to home directories causing the service role to not start.

The root cause of this issue is, when accessing Amazon cloud resources, Hue uses the AWS Boto SDK library. This AWS Boto library has a bug that restricts permissions in certain AWS regions in such a way that it provides access to users who should not have it, regardless of RAZ settings. This issue only affects users in specific AWS regions, listed below, and it does not affect all AWS customers.

Knowledge article

For the latest update on this issue see the corresponding Knowledge Article: TSB 2024-723: Hue Raz is using logger role to Read and Upload/Delete (write) files.

Known issues before 7.2.16

CDPD-58978: Batch query execution using Hue fails with Kerberos error

When you run Impala queries in a batch mode, you enounter failures with a Kerberos error even if the keytab is configured correctly. This is because submitting Impala, Sqoop, Pig, or pyspark queries in a batch mode launches a shell script Oozie job from Hue and this is not supported on a secure cluster.

There is no workaround. You can submit the queries individually.

Unable to delete, move, or rename directories within the S3 bucket from Hue

You may not be able to rename, move, or delete directories within your S3 bucket from the Hue web interface. This is because of an underlying issue, which will be fixed in a future release.

You can move, rename, or delete a directory using the HDFS commands as follows:

1. SSH into your CDP environment host.

2. To delete a directory within your S3 bucket, run the following command:

```
hdfs dfs -rm -r [***COMPLETE-PATH-TO-S3-BUCKET***]/[***DIREC
TORY-NAME***]
```

3. To rename a folder, create a new directory and run the following command to move files from the source directory to the target directory:

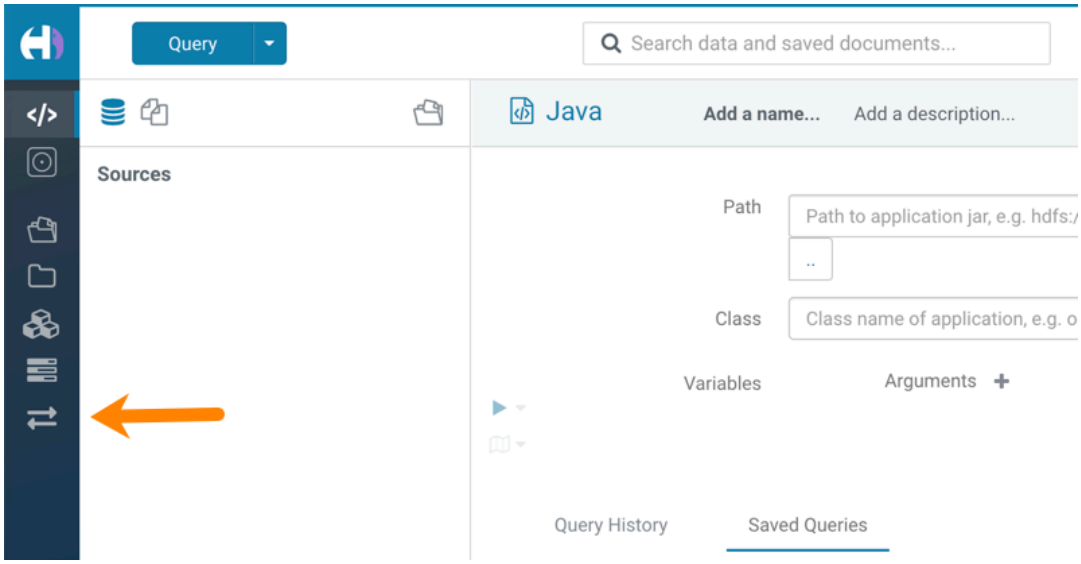
```
hdfs dfs -mkdir [***DIRECTORY-NAME***]
```

```
hdfs dfs -mv [***COMPLETE-PATH-TO-S3-BUCKET***]/[***SOURCE-D
IRECTORY***] [***COMPLETE-PATH-TO-S3-BUCKET***]/[***TARGET-D
IRECTORY***]
```

Hue Importer is not supported in the Data Engineering template

When you create a Data Hub cluster using the Data Engineering template, the Importer application is not supported in Hue.

Figure 1: Hue web UI showing Importer icon on the left assist panel



Unsupported features

CDPD-59595: Spark SQL does not work with all Livy servers that are configured for High Availability

SparkSQL support in Hue with Livy servers in HA mode is not supported. Hue does not automatically connect to one of the Livy servers. You must specify the Livy server in the Hue Advanced Configuration Snippet as follows:

```
[desktop]
[spark]
livy_server_url=http(s)://[***LIVY-FOR-SPARK3-SERVER-HOST***]:
[***LIVY-FOR-SPARK3-SERVER-PORT***]
```

Moreover, you may see the following error in Hue when you submit a SparkSQL query: Expecting value: line 2 column 1 (char 1). This happens when the Livy server does not respond to the request from Hue.

Specify all different Livy servers in the livy_server_url property one at a time and use the one which does not cause the issue.

Importing and exporting Oozie workflows across clusters and between different CDH versions is not supported

You can export Oozie workflows, schedules, and bundles from Hue and import them only within the same cluster if the cluster is unchanged. You can migrate bundle and coordinator jobs with their workflows only if their arguments have not changed between the old and the new cluster. For example, hostnames, NameNode, Resource Manager names, YARN queue names, and all the other parameters defined in the workflow.xml and job.properties files.

Using the import-export feature to migrate data between clusters is not recommended. To migrate data between different versions of CDH, for example, from CDH 5 to CDP 7, you must take the dump of the Hue database on the old cluster, restore it on the new cluster, and set up the database in the new environment. Also, the authentication method on the old and the new cluster should be the same because the Oozie workflows are tied to a user ID, and the exact user ID needs to be present in the new environment so that when a user logs into Hue, they can access their respective workflows.

Note: Migrating Oozie workflows from HDP clusters is not supported.

INSIGHT-3707: Query history displays "Result Expired" message

You see the "Result Expired" message under the Query History column on the **Queries** tab for queries which were run back to back. This is a known behaviour.

None.

Known Issues Iceberg

Learn about the known issues in Iceberg, the impact or changes to the functionality, and the workaround.

CDPD-55243: Spark can read stale data under certain conditions

Using inconsistent cases for database and table names of Iceberg tables in queries can lead to Spark reading stale data. In particular, after any write to the table (append, update, delete) in the same session, an old snapshot may still be read, unless all queries use consistent casing for database and table names.

Use consistent casing for database and table names in all queries.

CDPD-84220: Cannot query Iceberg tables

You cannot query existing Iceberg tables after you enable HDFS HA. This is because Iceberg stores the table path in the manifest files differently depending on whether the HDFS HA is enabled or not. After you enable HDFS HA, you might not be able to query the tables created prior to you enabling HDFS HA.

None.

Known Issues in Apache Impala

Learn about the known issues in Impala, the impact or changes to the functionality, and the workaround.

Impala known limitation when querying compacted tables

When the compaction process deletes the files for a table from the underlying HDFS location, the Impala service does not detect the changes as the compactions does not allocate new write ids. When the same table is queried from Impala it throws a 'File does not exist' exception that looks something like this:

```
Query Status: Disk I/O error on <node>:22000: Failed to open HDF
S file hdfs://nameservice1/warehouse/tablespace/managed/hive/<da
tabase>/<table>/xxxxx
Error(2): No such file or directory Root cause: RemoteException:
 File does not exist: /warehouse/tablespace/managed/hive/<data
base>/<table>/xxxx
```

Use the REFRESH/INVALIDATE statements on the affected table to overcome the 'File does not exist' exception.

HADOOP-15720: Queries stuck on failed HDFS calls and not timing out

In Impala 3.2 and higher, if the following error appears multiple times in a short duration while running a query, it would mean that the connection between the impalad and the HDFS NameNode is in a bad state.

```
"hdfsOpenFile() for <filename> at backend <hostname:port> failed
 to finish before the <hdfs_operation_timeout_sec> second timeout
 "
```

In Impala 3.1 and lower, the same issue would cause Impala to wait for a long time or not respond without showing the above error message.

Restart the impalad.

IMPALA-532: Impala should tolerate bad locale settings

If the LC_* environment variables specify an unsupported locale, Impala does not start.

Add LC_ALL="C" to the environment settings for both the Impala daemon and the Statestore daemon.

IMPALA-5605: Configuration to prevent crashes caused by thread resource limits

Impala could encounter a serious error due to resource usage under very high concurrency. The error message is similar to:

```
F0629 08:20:02.956413 29088 llvm-codegen.cc:111] LLVM hit fatal
 error: Unable to allocate section memory!
terminate called after throwing an instance of 'boost::exception_
detail::clone_impl<boost::exception_detail::error_info_injector<
boost::thread_resource_error> >'
```

To prevent such errors, configure each host running an `impalad` daemon with the following settings:

```
echo 2000000 > /proc/sys/kernel/threads-max
echo 2000000 > /proc/sys/kernel/pid_max
echo 8000000 > /proc/sys/vm/max_map_count
```

Add the following lines in /etc/security/limits.conf:

```
impala soft nproc 262144
impala hard nproc 262144
```

IMPALA-635: Avro Scanner fails to parse some schemas

The default value in Avro schema must match type of first union type, e.g. if the default value is null, then the first type in the UNION must be "null".

Swap the order of the fields in the schema specification. For example, use ["null", "string"] instead of ["string", "null"]. Note that the files written with the problematic schema must be rewritten with the new schema because Avro files have embedded schemas.

IMPALA-691: Process mem limit does not account for the JVM's memory usage

Some memory allocated by the JVM used internally by Impala is not counted against the memory limit for the impalad daemon.

To monitor overall memory usage, use the top command, or add the memory figures in the Impala web UI /memz tab to JVM memory usage shown on the /metrics tab.

IMPALA-9350: Ranger audit logs for applying column masking policies missing

Impala is not producing these logs.

None

IMPALA-1024: Impala BE cannot parse Avro schema that contains a trailing semi-colon

If an Avro table has a schema definition with a trailing semicolon, Impala encounters an error when the table is queried.

Remove trailing semicolon from the Avro schema.

IMPALA-1652: Incorrect results with basic predicate on CHAR typed column

When comparing a CHAR column value to a string literal, the literal value is not blank-padded and so the comparison might fail when it should match.

Use the RPAD() function to blank-pad literals compared with CHAR columns to the expected length.

IMPALA-1792: ImpalaODBC: Can not get the value in the SQLGetData(m-x th column) after the SQLBindCol(m th column)

If the ODBC SQLGetData is called on a series of columns, the function calls must follow the same order as the columns. For example, if data is fetched from column 2 then column 1, the SQLGetDa ta call for column 1 returns NULL.

Fetch columns in the same order they are defined in the table.

IMPALA-1821: Casting scenarios with invalid/inconsistent results

Using a CAST() function to convert large literal values to smaller types, or to convert special values such as NaN or Inf, produces values not consistent with other database systems. This could lead to unexpected results from queries.

IMPALA-2005: A failed CTAS does not drop the table if the insert fails

If a CREATE TABLE AS SELECT operation successfully creates the target table but an error occurs while querying the source table or copying the data, the new table is left behind rather than being dropped.

Drop the new table manually after a failed CREATE TABLE AS SELECT

IMPALA-2422: % escaping does not work correctly when occurs at the end in a LIKE clause

If the final character in the RHS argument of a LIKE operator is an escaped \% character, it does not match a % final character of the LHS argument.

IMPALA-2603: Crash: impala::Coordinator::ValidateCollectionSlots

A query could encounter a serious error if includes multiple nested levels of INNER JOIN clauses involving subqueries.

IMPALA-3094: Incorrect result due to constant evaluation in query with outer join

An OUTER JOIN query could omit some expected result rows due to a constant such as FALSE in another join clause. For example:

```
explain SELECT 1 FROM alltypestiny a1
  INNER JOIN alltypesagg a2 ON a1.smallint_col = a2.year AND fals
e
  RIGHT JOIN alltypes a3 ON a1.year = a1.bigint_col;
+-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\
-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-+
```

```
| Explain String                                          |
+-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\
-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-+
| Estimated Per-Host Requirements: Memory=1.00KB VCores=1 |
|                                                         |
| 00:EMPTYSET                                             |
+-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\
-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-\-+
```

IMPALA-3509: Breakpad minidumps can be very large when the thread count is high

The size of the breakpad minidump files grows linearly with the number of threads. By default, each thread adds 8 KB to the minidump size. Minidump files could consume significant disk space when the daemons have a high number of threads.

Add -\-minidump_size_limit_hint_kb=size to set a soft upper limit on the size of each minidump file. If the minidump file would exceed that limit, Impala reduces the amount of information for each thread from 8 KB to 2 KB. (Full thread information is captured for the first 20 threads, then 2 KB per thread after that.) The minidump file can still grow larger than the "hinted" size. For example, if you have 10,000 threads, the minidump file can be more than 20 MB.

IMPALA-4978: Impala requires FQDN from hostname command on Kerberized clusters

The method Impala uses to retrieve the host name while constructing the Kerberos principal is the gethostname() system call. This function might not always return the fully qualified domain name, depending on the network configuration. If the daemons cannot determine the FQDN, Impala does not start on a Kerberized cluster.

Test if a host is affected by checking whether the output of the `hostname` command includes the FQDN. On hosts where `hostname`, only returns the short name, pass the command-line flag ##hostname=*FULLY_QUALIFIED_DOMAIN_NAME* in the startup options of all Impala-related daemons.

IMPALA-6671: Metadata operations block read-only operations on unrelated tables

Metadata operations that change the state of a table, like COMPUTE STATS or ALTER RE COVER PARTITIONS, may delay metadata propagation of unrelated unloaded tables triggered by statements like DESCRIBE or SELECT queries.

Workaround: None

IMPALA-7072: Impala does not support Heimdal Kerberos

CDPD-28139: Set spark.hadoop.hive.stats.autogather to false by default

As an Impala user, if you submit a query against a table containing data ingested using Spark and you are concerned about the quality of the query plan, you must run COMPUTE STATS against such a table in any case after an ETL operation because numRows created by Spark could be incorrect. Also, use other stats computed by COMPUTE STATS, e.g., Number of Distinct Values (NDV) and NULL count for good selectivity estimates.

For example, when a user ingests data from a file into a partition of an existing table using Spark, if spark.hadoop.hive.stats.autogather is not set to false explicitly, numRows associated with this partition would be 0 even though there is at least one row in the file. To avoid this, the workaround is to set "spark.hadoop.hive.stats.autogather=false" in the "Spark Client Advanced Configuration Snippet (Safety Valve) for spark-conf/spark-defaults.conf" in Spark's CM Configuration section.

Known Issues in Apache Kafka

Learn about the known issues in Apache Kafka, the impact or changes to the functionality, and the workaround.

Known Issues

OPSAPS-59553: SMM's bootstrap server config should be updated based on Kafka's listeners

SMM does not show any metrics for Kafka or Kafka Connect when multiple listeners are set in Kafka.

Workaround: SMM cannot identify multiple listeners and still points to bootstrap server using the default broker port (9093 for SASL_SSL). You would have to override bootstrap server URL (hostname:port as set in the listeners for broker) in the following path:

Cloudera Manager > SMM > Configuration > Streams Messaging Manager Rest Admin Server Advanced Configuration Snippet (Safety Valve) for streams-messaging-manager.yaml > Save Changes > Restart SMM.

The offsets.topic.replication.factor property must be less than or equal to the number of live brokers

The offsets.topic.replication.factor broker configuration is now enforced upon auto topic creation. Internal auto topic creation will fail with a GROUP_COORDINATOR_NOT_AVAILABLE error until the cluster size meets this replication factor requirement.

None

Requests fail when sending to a nonexistent topic with auto.create.topics.enable set to true

The first few produce requests fail when sending to a nonexistent topic with auto.create.topics.enable set to true.

Increase the number of retries in the producer configuration setting retries.

KAFKA-2561: Performance degradation when SSL Is enabled

In some configuration scenarios, significant performance degradation can occur when SSL is enabled. The impact varies depending on your CPU, JVM version, Kafka configuration, and message size. Consumers are typically more affected than producers.

Configure brokers and clients with ssl.secure.random.implementation = SHA1PRNG. It often reduces this degradation drastically, but its effect is CPU and JVM dependent.

OPSAPS-43236: Kafka garbage collection logs are written to the process directory

By default Kafka garbage collection logs are written to the agent process directory. Changing the default path for these log files is currently unsupported.

None

OPSAPS-65485: Selecting the Require Connectors To Override Kafka Client JAAS Configuration property causes automatic Kafka Connect startup retries to fail

If the Require Connectors To Override Kafka Client JAAS Configuration property is selected for the Kafka Connect role and the role fails to start due to any reason, all automatic retries to start the role will also fail. This is true even if the root cause of the initial startup failure was caused by an intermittent issue.

Review the Kafka connect logs and resolve the root cause of the startup failure. Once the issue is resolved, restart the Kafka Connect role manually. Alternatively, clear the Require Connectors To Override Kafka Client JAAS Configuration property.

RANGER-3809: Idempotent Kafka producer fails to initialize due to an authorization failure

Kafka producers that have idempotence enabled require the Idempotent Write permission to be set on the cluster resource in Ranger. If permission is not given, the client fails to initialize and an error similar to the following is thrown:

```
org.apache.kafka.common.KafkaException: Cannot execute transacti
onal method because we are in an error state
    at org.apache.kafka.clients.producer.internals.TransactionM
anager.maybeFailWithError(TransactionManager.java:1125)
    at org.apache.kafka.clients.producer.internals.TransactionMa
nager.maybeAddPartition(TransactionManager.java:442)
```

```
    at org.apache.kafka.clients.producer.KafkaProducer.doSend(K
afkaProducer.java:1000)
    at org.apache.kafka.clients.producer.KafkaProducer.send(Kafk
aProducer.java:914)
    at org.apache.kafka.clients.producer.KafkaProducer.send(Kafk
aProducer.java:800)
    .
    .
    .
Caused by: org.apache.kafka.common.errors.ClusterAuthorization
Exception: Cluster authorization failed.
```

Idempotence is enabled by default for clients in Kafka 3.0.1, 3.1.1, and any version after 3.1.1. This means that any client updated to 3.0.1, 3.1.1, or any version after 3.1.1 is affected by this issue.

This issue has two workarounds, do either of the following:

- Explicitly disable idempotence for the producers. This can be done by setting enable.idempoten ce to false.
- Update your policies in Ranger and ensure that producers have Idempotent Write permission on the cluster resource.

CDPD-29307: Kafka producer entity stays in incomplete state in Atlas

Atlas creates incomplete Kafka client entities that are postfixed with the metadata namespace.

None

DBZ-4990: The Debezium Db2 Source connector does not support schema evolution

The Debezium Db2 Source connector does not support the evolution (updates) of schemas. In addition, schema change events are not emitted to the schema change topic if there is a change in the schema of a table that is in capture mode. For more information, see DBZ-4990.

None.

CDPD-53179: Amazon S3 sink connector fails when buffer size is reached

If there is more than 5 MB (buffer size) of data available in a Kafka source topic and the connector receives more than 5 MB of data in a single poll, the connector tries to upload all the data as a multipart upload to S3. The upload, however, fails.

Decrease the Offset Flush Interval Kafka service property. Decreasing the value of this property increases how frequently connectors commit data. If the connectors commit more frequently, each commit will contain less data. Cloudera advises caution if you decide to change the value of this property because:

- The interval you configure is applied to all Kafka Connect connectors.
- Decreasing the interval might result in an increased number of files created in S3.

CDPD-48822: AvroConverter ignores default values when converting from Avro to Connect schema

AvroConverter does not propagate field default values when converting Avro schemas to Connect schemas.

None

Unsupported Features

The following Kafka features are not supported in Cloudera Data Platform:

- Only Java and .Net based clients are supported. Clients developed with C, C++, Python, and other languages are currently not supported.
- The Kafka default authorizer is not supported. This includes setting ACLs and all related APIs, broker functionality, and command-line tools.
- SASL/SCRAM is only supported for delegation token based authentication. It is not supported as a standalone authentication mechanism.

- The Cloudera `AvroConverter` (`com.cloudera.dim.kafka.connect.converts.AvroConverter`) is not supported with the Debezium Kafka Connect connectors shipped in Cloudera Runtime.

Limitations

Collection of Partition Level Metrics May Cause Cloudera Manager's Performance to Degrade

If the Kafka service operates with a large number of partitions, collection of partition level metrics may cause Cloudera Manager's performance to degrade.

If you are observing performance degradation and your cluster is operating with a high number of partitions, you can choose to disable the collection of partition level metrics.

Important: If you are using SMM to monitor Kafka or Cruise Control for rebalancing Kafka partitions, be aware that both SMM and Cruise Control rely on partition level metrics. If partition level metric collection is disabled, SMM will not be able to display information about partitions. In addition, Cruise Control will not operate properly.

Complete the following steps to turn off the collection of partition level metrics:

1. Obtain the Kafka service name:
   a. In Cloudera Manager, Select the Kafka service.
   b. Select any available chart, and select Open in Chart Builder from the configuration icon drop-down.
   c. Find $SERVICENAME= near the top of the display.

      The Kafka service name is the value of $SERVICENAME.
2. Turn off the collection of partition level metrics:
   a. Go to HostsHosts Configuration.
   b. Find and configure the Cloudera Manager Agent Monitoring Advanced Configuration Snippet (Safety Valve) configuration property.

      Enter the following to turn off the collection of partition level metrics:

      ```
      [KAFKA_SERVICE_NAME]_feature_send_broker_topic_partition_ent
      ity_update_enabled=false
      ```

      Replace [KAFKA_SERVICE_NAME] with the service name of Kafka obtained in step 1. The service name should always be in lower case.
   c. Click Save Changes.

Known Issues in Apache Knox

Learn about the known issues in Knox, the impact or changes to the functionality, and the workaround.

CDPD-3125: Logging out of Atlas does not manage the external authentication

At this time, Atlas does not communicate a log-out event with the external authentication management, Apache Knox. When you log out of Atlas, you can still open the instance of Atlas from the same web browser without re-authentication.

To prevent additional access to Atlas, close all browser windows and exit the browser.

Known Issues in Apache Kudu

Learn about the known issues in Kudu, the impact or changes to the functionality, and the workaround.

Kudu supports only coarse-grain authorization. Kudu does not yet support integration with Atlas.

None

Kudu HMS Sync is disabled and is not yet supported

None

Kerberos authentication fails with rdns disabled

When rdns is set to false, the Kudu Java client does not retain the original hostname, and replaces them with the resolved IP addresses. This prevents Kerberos authentication from working properly. For more information, see KUDU-3415.

Add a "ranger_kudu_plugin_service_name" configuration to the Kudu Master configuration group in the blueprint with the value "{{GENERATED_RANGER_SERVICE_NAME}}"

None

Known Issues in Apache Oozie

Learn about the known issues in Oozie, the impact or changes to the functionality, and the workaround.

CDPD-41274: HWC + Oozie issue: Could not open client transport with JDBC Uri

Currently only Spark cluster mode is supported in the Oozie Spark Action with Hive Warehouse Connector (HWC).

Use Spark action in cluster mode.

```
Use Spark action in cluster mode.
                        <spark xmlns="uri:oozie:spark-action:1
.0">
                        ...
                        <mode>cluster</mode>
                        ...
                        </spark>
```

CDPD-26975: Using the ABFS / S3A connectors in an Oozie workflow where the operations are "secured" may trigger an IllegalArgumentException with the error message java.net.URISyntaxException: Relative path in absolute URI.

Set the following XML configuration in the Datahub cluster's Cloudera Manager:

1. In the Cloudera Manager Admin Console, go to the Oozie service.
2. Click the Configuration tab.
3. In the Oozie Server Advanced Configuration Snippet (Safety Valve) for oozie-site.xml field, set the following:

   Set the following if you are using Amazon S3:

   ```
    <property>
         <name>oozie.service.HadoopAccessorService.fs.s3a</name>
         <value>fs.s3a.buffer.dir=/tmp/s3a</value>
      </property>
   ```

   Set the following if you are using ABFS:

   ```
   <property>
         <name>oozie.service.HadoopAccessorService.fs.abfs</name>
         <value>fs.azure.buffer.dir=/tmp/abfs</value>
      </property>
      <property>
         <name>oozie.service.HadoopAccessorService.fs.abfss</name
   >
         <value>fs.azure.buffer.dir=/tmp/abfss</value>
      </property>
   ```

4. Enter a Reason for change, and then click Save Change to commit the changes.
5. Restart the Oozie service.

Oozie jobs fail (gracefully) on secure YARN clusters when JobHistory server is down

If the JobHistory server is down on a YARN (MRv2) cluster, Oozie attempts to submit a job, by default, three times. If the job fails, Oozie automatically puts the workflow in a SUSPEND state.

When the JobHistory server is running again, use the resume command to inform Oozie to continue the workflow from the point at which it left off.

Unsupported Feature

The following Oozie features are currently not supported in Cloudera Data Platform:

- Non-support for Pig action (CDPD-1070)
- Conditional coordinator input logic

Cloudera does not support using Derby database with Oozie. You can use it for testing or debugging purposes, but Cloudera does not recommend using it in production environments. This could cause failures while upgrading from CDH to CDP.

BUG-123856: Upgrade fails while configuring Oozie server.

None

Known Issues in Apache Phoenix

There are no known issues for Phoenix in Cloudera Runtime 7.2.16.

Known Issues in Apache Ranger

Learn about the known issues in Ranger, the impact or changes to the functionality, and the workaround.

CDPD-3296: Audit files for Ranger plugin components do not appear immediately in S3 after cluster creation

For Ranger plugin components (Atlas, Hive, HBase, etc.), audit data is updated when the applicable audit file is rolled over. The default Ranger audit rollover time is 24 hours, so audit data appears 24 hours after cluster creation.

To see the audit logs in S3 before the default rollover time of 24 hours, use the following steps to override the default value in the Cloudera Manager safety valve for the applicable service.

1. On the Configuration tab in the applicable service, select Advanced under CATEGORY.
2. Click the + icon for the <service_name> Advanced Configuration Snippet (Safety Valve) for ranger-<service_name>-audit.xml property.
3. Enter the following property in the Name box:

   xasecure.audit.destination.hdfs.file.rollover.sec.
4. Enter the desired rollover interval (in seconds) in the Value box. For example, if you specify 180, the audit log data is updated every 3 minutes.
5. Click Save Changes and restart the service.

Note: You can also use xasecure.audit.destination.hdfs.file.rollover.period parameter to override the default rollover time of 24 hours. The difference is that when xasecure .audit.destination.hdfs.file.rollover.period is set, it will be closing the file by absolute time.

Example - If you configure 1 day, exactly at 23.59.59 of that day, the file gets closed. Where as with xasecure.audit.destination.hdfs.file.rollover.sec, the 1 day is related to when the process is started.

CDPD-12644: Ranger Key Names cannot be reused with the Ranger KMS KTS service

Key names cannot be reused with the Ranger KMS KTS service. If the key name of a delete key is reused, the new key can be successfully created and used to create an encryption zone, but data cannot be written to that encryption zone.

Use only unique key names when creating keys.

OPSAPS-70387: The DataHub cluster deletion process does not delete the Ranger entries which created for the same cluster.

If the user wants to create a new DataHub cluster with same old name then it fails because as there was an entry with the same name already in Ranger.

User must delete the Ranger entries manually which contains the DataHub cluster name.

CDPD-17962: Ranger roles do not work when you upgrade from any CDP Private Cloud Base to CDP Private cloud base. Roles which are created prior to upgrade work as expected, issue is only for new roles created post upgrade and authorization enforced via ranger policies wont work for these new roles. This behavior is only observed with the upgraded cluster; a newly installed cluster does not show this behavior.

There are two possible workarounds to resolve this issue:

1. Update database entries (Recommended):
   - select * from x_ranger_global_state where state_name='RangerRole';
   - update x_ranger_global_state set app_data='{"Version":"2"}' where state_name='RangerRole';

   Or
2. Add a property in safety valve under ranger-admin-site which will bypass the getAppDataVersion method:

Technical Service Bulletins

TSB 2023-655: Apache Ranger (Ranger) S3 policies for READ or WRITE are not evaluated on RAZ-enabled CDP Public Cloud 7.2.16 environments

Fine-grained Ranger Authorization Service (RAZ) enables users to define Ranger policies on S3 (cm_s3) paths. Due to a recent change in Cloudera Data Platform (CDP) Public Cloud 7.2.16.0, the access type is processed incorrectly. This results in read-only or write-only Ranger policies not being evaluated. Ranger policies with all, or both read and write permissions are not affected.

For the latest update on this issue see the corresponding Knowledge article: TSB 2023-655: Apache Ranger (Ranger) S3 policies for READ or WRITE are not evaluated on RAZ-enabled CDP Public Cloud 7.2.16 environments

Known Issues in Schema Registry

Learn about the known issues in Schema Registry, the impact or changes to the functionality, and the workaround.

CDPD-48568: JAR storage does not work on AWS S3 for Schema Registry

If you are using AWS S3 to store the serializer and deserializer JAR files used by Schema Registry, the Schema Registry server might not be able reach these JAR files. As a result, message serialization and deserialization will not work with the affected schemas. If you are affected by this issue, class not found errors related to AWS library classes will be present in the Schema Registry server logs.

Ensure that the aws-java-sdk-bundle-*[***VERSION***]*.jar file is available in /opt/cloudera/parcels/CDH/lib/schemaregistry/hadoop-plugin/hadoop-schema-registry-plugin-impl/. This is the location that contains all the JAR files for the Hadoop plugin classpath.

1. Find the aws-java-sdk-bundle-*[***VERSION***]*.jar file.

   The file is located in the folder that contains all common JAR files for parcels. You can find it using the following command:

   ```
   find /opt/cloudera/parcels/CDH/jars -iname '*aws-java-sdk-bu
   ndle*.jar
   ```

   The output of the command lists multiple versions. Take note of the latest version. For example:

   ```
   /opt/cloudera/parcels/CDH/jars/aws-java-sdk-bundle-1.12.316.
   jar
   ```

2. Create a symlink in /opt/cloudera/parcels/CDH/lib/schemaregistry/hadoop-plugin/hadoop-schema-registry-plugin-impl/ that points to the aws-java-sdk-bundle-*[***VERSION***]*.jar. For example:

   ```
   ln -s /opt/cloudera/parcels/CDH/jars/aws-java-sdk-bundle-1.1
   2.316.jar /opt/cloudera/parcels/CDH/lib/schemaregistry/hadoop-
   plugin/hadoop-schema-registry-plugin-impl/aws-java-sdk-bundle-
   1.12.316.jar
   ```

3. Restart the Schema Registry service.

CDPD-54379: KafkaJsonSerializer and KafkaJsonDeserializer do not allow null values

`KafkaJsonSerializer` and `KafkaJsonDeserializer` do not allow the data to be null, resulting in a `NullPointerException` (NPE).

None.

CDPD-49217 and CDPD-50309: Schema Registry caches user group membership indefinitely

Schema Registry caches the Kerberos user and group information indefinitely and does not catch up on group membership changes.

Restart Schema Registry after group membership changes.

CDPD-56890: New schemas cannot be created following an upgrade

If you delete the latest version of a schema (the one with the highest ID) from the Schema Registry database before an upgrade, you might not be able to create new schemas after you upgrade the cluster to a newer version.

Important: In CDP Public Cloud, this issue only manifests when upgrading from Cloudera Runtime 7.2.12 or lower to 7.2.14 or higher.

1. Access the Schema Registry database. Go to Cloudera Manager Schema Registry Configuration and search for "database" if you don't know the name, host, or port of the Schema Registry database.
2. Cross reference the ID's in the schemaVersionId column of the schmema_version_state table with the ID's found in the schema_version_info table.
3. Delete all records from the schema_version_state table that contains a schemaVersionId not present in the schema_version_info table.

CDPD-58265: Schema Registry Client incorrectly applies SSL configuration

The Cloudera distributed Schema Registry Java client might fail to apply the SSL configurations correctly with concurrent access in Jersey clients due to a Jersey issue related to JDK.

Before using HttpsURLConnection in any form concurrently, call javax.net.ssl.HttpsURLConnection.getDefaultSSLSocketFactory() once in the custom client application.

CDPD-48822: AvroConverter ignores default values when converting from Avro to Connect schema

AvroConverter does not propagate field default values when converting Avro schemas to Connect schemas.

None

CDPD-55381: Schema Registry issues authentication cookie for the authorized user, not for the authenticated one

When the authenticated user is different from the authorized user, which can happen when Schema Registry is used behind Knox, authorization issues can occur for subsequent requests as the authentication cookie in Schema Registry stores the authorized user.

Access Schema Registry directly, without using Knox, if possible. If not, ensure that the name of the end user that tries to connect does not begin with knox.

CDPD-60160: Schema Registry Atlas integration does not work with Oracle databases

Schema Registry is unable to create entities in Atlas if Schema Registry uses an Oracle database. The following will be present in the Schema Registry log if you are affected by this issue:

```
ERROR com.cloudera.dim.atlas.events.AtlasEventsProcessor: An err
or occurred while processing Atlas events.
java.lang.IllegalArgumentException: Cannot invoke com.hortonworks
.registries.schemaregistry.AtlasEventStorable.setType on bean cl
ass 'class com.hortonworks.registries.schemaregistry.AtlasEventS
torable' - argument type mismatch - had objects of type "java.la
ng.Long" but expected signature "java.lang.Integer"
```

This issue causes the loss of audit data on Oracle environments.

None.

CDPD-48853: Schemas created with the Confluent Schema Registry API cannot be viewed in the UI

Schemas created in Cloudera Schema Registry using the Confluent Schema Registry API are not visible in the Cloudera Schema Registry UI.

In addition, the `/api/v1/schemaregistry/search/schemas/aggregated` endpoint of the Cloudera Schema Registry API does not return schemas created with the Confluent Schema Registry API.

A typical case where this issue can manifest is when you are using the Confluent Avro converter for SerDes in a Kafka Connect connector and the connector connects to Cloudera Schema Registry. That is, the key.converter and/or value.converter properties of the connector are set to `io.confluent.connect.avro.AvroConverter`, and key.converter.schema.registry.url and/or value.converter.schema.registry.url are set to a Cloudera Schema Registry server URL.

None.

CDPD-58949: Schemas are de-duplicated on import

On import, Schema Registry de-duplicates schema versions based on their fingerprints. This means that schemas which are considered functionally equivalent in SR get de-duplicated. As a result, some schema versions are not created, and their IDs do not become valid IDs in SR.

None.

CDPD-58990: getSortedSchemaVersions method orders by schemaVersionId instead of version number

On validation, Schema Registry orders schema versions based on ID instead of version number. In some situations, this can cause validation with the LATEST level to compare the new schema version to a non-latest version.

This situation can occur when an older version of a schema has a higher ID than the newer version of a schema, for example, when the older version is imported with an explicit ID.

None.

Known Issues in Cloudera Search

Learn about the known issues in Cloudera Search, the impact or changes to the functionality, and the workaround.

Note: From Cloudera Runtime 7.2.15 and higher, Cloudera Search known issues are listed under Apache Solr.

Known Issues in Apache Solr

Learn about the known issues in Solr, the impact or changes to the functionality, and the workaround.

Known Issues

Changing the default value of Client Connection Registry HBase configuration parameter causes HBase MRIT job to fail

If the value of the HBase configuration property Client Connection Registry is changed from the default ZooKeeper Quorum to Master Registry then the Yarn job started by HBase MRIT fails with a similar error message:

```
Caused by: org.apache.hadoop.hbase.exceptions.MasterRegistryFetc
hException: Exception making rpc to masters [quasar-bmyccr-2.qua
sar-bmyccr.root.hwx.site,22001,-1]
        at org.apache.hadoop.hbase.client.MasterRegistry.lambda$g
roupCall$1(MasterRegistry.java:244)
        at org.apache.hadoop.hbase.util.FutureUtils.lambda$addLi
stener$0(FutureUtils.java:68)
        at java.util.concurrent.CompletableFuture.uniWhenCompl
ete(CompletableFuture.java:774)
        at java.util.concurrent.CompletableFuture.uniWhenComplet
eStage(CompletableFuture.java:792)
        at java.util.concurrent.CompletableFuture.whenComplete(Co
mpletableFuture.java:2153)
        at org.apache.hadoop.hbase.util.FutureUtils.addListener(F
utureUtils.java:61)
        at org.apache.hadoop.hbase.client.MasterRegistry.groupCa
ll(MasterRegistry.java:228)
        at org.apache.hadoop.hbase.client.MasterRegistry.call(Ma
sterRegistry.java:265)
        at org.apache.hadoop.hbase.client.MasterRegistry.getMetaR
egionLocations(MasterRegistry.java:282)
        at org.apache.hadoop.hbase.client.ConnectionImplementati
on.locateMeta(ConnectionImplementation.java:900)
        at org.apache.hadoop.hbase.client.ConnectionImplementat
ion.locateRegion(ConnectionImplementation.java:867)
        at org.apache.hadoop.hbase.client.ConnectionImplementati
on.relocateRegion(ConnectionImplementation.java:850)
        at org.apache.hadoop.hbase.client.ConnectionImplementat
ion.locateRegionInMeta(ConnectionImplementation.java:981)
        at org.apache.hadoop.hbase.client.ConnectionImplementa
tion.locateRegion(ConnectionImplementation.java:870)
        at org.apache.hadoop.hbase.client.RpcRetryingCallerWith
ReadReplicas.getRegionLocations(RpcRetryingCallerWithReadReplica
s.java:319)
        ... 21 more
Caused by: org.apache.hadoop.hbase.client.RetriesExhaustedExcept
ion: Failed contacting masters after 1 attempts.
Exceptions:
java.io.IOException: Call to address=quasar-bmyccr-2.quasar-bmy
ccr.root.hwx.site/172.27.19.4:22001 failed on local exception: j
```

```
ava.io.IOException: java.lang.RuntimeException: Found no valid a
uthentication method from options
        at org.apache.hadoop.hbase.client.MasterRegistry.lambda
$groupCall$1(MasterRegistry.java:243)
        ... 35 more
```

Add the following line to the MRIT command line:

```
-D 'hbase.client.registry.impl=org.apache.hadoop.hbase.client.ZK
ConnectionRegistry'
```

Apache Tika upgrade may break morphlines indexing

The upgrade of Apache Tika from 1.27 to 2.3.0 brought potentially breaking changes for morphlines indexing. Duplicate/triplicate keys names were removed and certain parser class names were changed (For example, org.apache.tika.parser.jpeg.JpegParser changed to org.apache.tika.parser.image.JpegParser).

To avoid morphline commands failing after the upgrade, do the following:

- Check if key name changes affect your morphlines. For more information, see *Removed duplicate/triplicate keys* in Migrating to Tika 2.0.0.
- Check if the name of any parser you use has changed. For more information, see the Apache Tika API documentation.

Update your morphlines if necessary.

CDPD-28432: HBase Lily indexer REST port does not support SSL

When using the --http argument for the hbase-indexer command line tool to invoke Lily indexer through REST API, you can add/list/remove indexers with any user without the need for authentication.

Switch off the REST API setting the hbaseindexer.httpserver.disabled environment parameter to true (by default this is false). This switches off the REST interface, so noone can use the --http argument when using the hbase-indexer command line tool. This also means that users need to authenticate as hbase user in order to use the hbase-indexer tool.

CDH-77598: Indexing fails with socketTimeout

Starting from CDH 6.0, the HTTP client library used by Solr has a default socket timeout of 10 minutes. Because of this, if a single request sent from an indexer executor to Solr takes more than 10 minutes to be serviced, the indexing process fails with a timeout error.

This timeout has been raised to 24 hours. Nevertheless, there still may be use cases where even this extended timeout period proves insufficient.

If your MapreduceIndexerTool or HBaseMapreduceIndexerTool batch indexing jobs fail with a timeout error during the go-live (Live merge, MERGEINDEXES) phase (This means the merge takes longer than 24 hours).

Use the --go-live-timeout option where the timeout can be specified in milliseconds.

CDPD-12450: CrunchIndexerTool Indexing fails with socketTimeout

The http client library uses a socket timeout of 10 minutes. The Spark Crunch Indexer does not override this value, and in case a single batch takes more than 10 minutes, the entire indexing job fails. This can happen especially if the morphlines contain DeleteByQuery requests.

Try the following workarounds:

- Check the batch size of your indexing job. Sending too large batches to Solr might increase the time needed on the Solr server to process the incoming batch.
- If your indexing job uses deleteByQuery requests, consider using deleteById wherever possible as deleteByQuery involves a complex locking mechanism on the Solr side which makes processing the requests slower.

- Check the number of executors for your Spark Crunch Indexer job. Too many executors can overload the Solr service. You can configure the number of executors by using the --mappers parameter
- Check that your Solr installation is correctly sized to accommodate the indexing load, making sure that the number of Solr servers and the number of shards in your target collection are adequate.
- The socket timeout for the connection can be configured in the morphline file. Add the solrClie ntSocketTimeout parameter to the solrLocator command

  Example

```
SOLR_LOCATOR :
{
  collection : test_collection
  zkHost : "zookeeper1.example.corp:2181/solr"
# 10 minutes in milliseconds
  solrClientSocketTimeout: 600000
  # Max number of documents to pass per RPC from morphline to
 Solr Server
  # batchSize : 10000
}
```

CDPD-29289: HBaseMapReduceIndexerTool fails with socketTimeout

The http client library uses a socket timeout of 10 minutes. The HBase Indexer does not override this value, and in case a single batch takes more than 10 minutes, the entire indexing job fails.

You can overwrite the default 600000 millisecond (10 minute) socket timeout in HBase indexer using the --solr-client-socket-timeout optional argument for the direct writing mode (when the value of the --reducers optional argument is set to 0 and mappers directly send the data to the live Solr).

CDPD-20577: Splitshard operation on HDFS index checks local filesystem and fails

When performing a shard split on an index that is stored on HDFS, SplitShardCmd still evaluates free disk space on the local file system of the server where Solr is installed. This may cause the command to fail, perceiving that there is no adequate disk space to perform the shard split.

Run the following command to skip the check for sufficient disk space altogether:

- On nonsecure clusters:

```
curl 'http://$[***SOLR_SERVER_HOSTNAME***]:8983/so
lr/admin/collections?action=SPLITSHARD&collectio
n=[***COLLECTION_NAME***]&shard=[***SHARD_TO_SPLIT***]&skipFre
eSpaceCheck=true'
```

- On secure clusters:

```
curl -k -u : --negotiate 'http://
$[***SOLR_SERVER_HOSTNAME***]:8985/solr/admin/collections
?action=SPLITSHARD&collection=[***COLLECTION_NAME***]&sha
rd=[***SHARD_TO_SPLIT***]&skipFreeSpaceCheck=true'
```

Replace *[***SOLR_SERVER_HOSTNAME***]* with a valid Solr server hostname, *[***COLLECTION_NAME***]* with the collection name, and *[***SHARD_TO_SPLIT***]* with the ID of the shard to split.

To verify that the command executed succesfully, check overseer logs for a similar entry:

```
2021-02-02 12:43:23.743 INFO  (OverseerThreadFactory-9-thread-5-
processing-n:myhost.example.com:8983_solr) [c:example s:shard1
  ] o.a.s.c.a.c.SplitShardCmd Skipping check for sufficient disk
 space
```

DOCS-5717: Lucene index handling limitation

The Lucene index can only be upgraded by one major version. Solr 8 will not open an index that was created with Solr 6 or earlier.

There is no workaround, you need to reindex collections.

CDH-22190: CrunchIndexerTool which includes Spark indexer requires specific input file format specifications

If the --input-file-format option is specified with CrunchIndexerTool, then its argument must be text, avro, or avroParquet, rather than a fully qualified class name.

None

CDH-26856: Field value class guessing and Automatic schema field addition are not supported with the MapReduceIndexerTool nor with the HBaseMapReduceIndexerTool.

The MapReduceIndexerTool and the HBaseMapReduceIndexerTool can be used with a Managed Schema created via NRT indexing of documents or via the Solr Schema API. However, neither tool supports adding fields automatically to the schema during ingest.

Define the schema before running the MapReduceIndexerTool or HBaseMapReduceIndexerTool. In non-schemaless mode, define in the schema using the schema.xml file. In schemaless mode, either define the schema using the Solr Schema API or index sample documents using NRT indexing before invoking the tools. In either case, Cloudera recommends that you verify that the schema is what you expect, using the List Fields API command.

CDH-19407: The Browse and Spell Request Handlers are not enabled in schemaless mode

The Browse and Spell Request Handlers require certain fields to be present in the schema. Since those fields cannot be guaranteed to exist in a Schemaless setup, the Browse and Spell Request Handlers are not enabled by default.

If you require the Browse and Spell Request Handlers, add them to the solrconfig.xml configuration file. Generate a non-schemaless configuration to see the usual settings and modify the required fields to fit your schema.

CDH-17978: Enabling blockcache writing may result in unusable indexes.

It is possible to create indexes with solr.hdfs.blockcache.write.enabled set to true. Such indexes may appear corrupt to readers, and reading these indexes may irrecoverably corrupt indexes. Blockcache writing is disabled by default.

None

CDH-58276: Users with insufficient Solr permissions may receive a "Page Loading" message from the Solr Web Admin UI.

Users who are not authorized to use the Solr Admin UI are not given a page explaining that access is denied to them, instead receive a web page that never finishes loading.

None

CDH-15441: Using MapReduceIndexerTool or HBaseMapReduceIndexerTool multiple times may produce duplicate entries in a collection.

Repeatedly running the MapReduceIndexerTool on the same set of input files can result in duplicate entries in the Solr collection. This occurs because the tool can only insert documents and cannot update or delete existing Solr documents. This issue does not apply to the HBaseMapReduceIndexerTool unless it is run with more than zero reducers.

To avoid this issue, use HBaseMapReduceIndexerTool with zero reducers. This must be done without Kerberos.

Note: This workaround is only valid for HBaseMapReduceIndexerTool. There is no workaround for MapReduceIndexerTool

CDH-58694: Deleting collections might fail if hosts are unavailable.

It is possible to delete a collection when hosts that host some of the collection are unavailable. After such a deletion, if the previously unavailable hosts are brought back online, the deleted collection may be restored.

Ensure all hosts are online before deleting collections.

Unsupported Features

The following Solr features are currently not supported in Cloudera Data Platform:

- Package Management System
- HTTP/2
- Solr SQL/JDBC
- Graph Traversal
- Cross Data Center Replication (CDCR)
- SolrCloud Autoscaling
- HDFS Federation
- Saving search results
- Solr contrib modules (Spark, MapReduce and Lily HBase indexers are not contrib modules but part of Cloudera's distribution of Solr itself, therefore they are supported).

Known Issues in Apache Spark

Learn about the known issues in Spark, the impact or changes to the functionality, and the workarounds.

CDPD-217: The Apache Spark connector is not supported

The old *Apache Spark - Apache HBase Connector* (shc) is not supported in CDP releases.

Use the new HBase-Spark connector shipped in CDP release.

CDPD-3038: Launching pyspark displays several HiveConf warning messages

When pyspark starts, several Hive configuration warning messages are displayed, similar to the following:

```
19/08/09 11:48:04 WARN conf.HiveConf: HiveConf of name hive.vect
orized.use.checked.expressions does not exist
19/08/09 11:48:04 WARN conf.HiveConf: HiveConf of name hive.te
z.cartesian-product.enabled does not exist
```

These errors can be safely ignored.

Known Issues for Apache Sqoop

Learn about the known issues in Sqoop, the impact or changes to the functionality, and the workaround.

Unable to read Sqoop metastore created by an older HSQLDB version

If you have upgraded to CDP Public Cloud 7.2.16 or higher versions, you may encounter issues in reading the Sqoop metastore that was created using an older version of HyperSQL Database (HSQLDB).

Cloudera upgraded the HSQLDB dependency from 1.8.0.10 to 2.7.1 and this causes incompatibility issues in Sqoop jobs that are stored in HSQLDB.

After upgrading to CDP Public Cloud 7.2.16, you must upgrade the Sqoop metastore and convert the database files to a format that can easily be read by HSQLDB 2.7.1. For more information, see Troubleshooting Apache Sqoop issues.

Using direct mode causes problems

Using direct mode has several drawbacks:

- Imports can cause intermittent an overlapping input split.
- Imports can generate duplicate data.
- Many problems, such as intermittent failures, can occur.
- Additional configuration is required.

Stop using direct mode. Do not use the --direct option in Sqoop import or export commands.

CDPD-3089: Avro, S3, and HCat do not work together properly

Importing an Avro file into S3 with HCat fails with Delegation Token not available.

Parquet columns inadvertently renamed

Column names that start with a number are renamed when you use the --as-parquetfile option to import data.

Prepend column names in Parquet tables with one or more letters or underscore characters.

Importing Parquet files might cause out-of-memory (OOM) errors

Importing multiple megabytes per row before initial-page-run check (ColumnWriter) can cause OOM. Also, rows that vary significantly by size so that the next-page-size check is based on small rows, and is set very high, followed by many large rows can also cause OOM.

None

Known Issues in Streams Messaging Manager

Learn about the known issues for Streams Messaging Manager in Cloudera Runtime 7.2.16.

CDPD-39313: Some numbers are not rendered properly in SMM UI

Very large numbers can be imprecisely represented on the UI. For example, bytes larger than 8 petabytes would lose precision.

None.

CDPD-46465: Searching for workers on the connector overview page freezes the page

Navigating to the SMM UI Connectors page selecting the Cluster Profile tab and filtering for hosts using the search field causes the page to freeze.

None.

CDPD-46728: SMM UI shows the consumerGroup instead of the instances on the Profile page's right hand side

On the ConsumerGroupDetail page, SMM UI shows the group instead of its instances on the right hand side table.

None.

OPSAPS-59553: SMM's bootstrap server config should be updated based on Kafka's listeners

SMM does not show any metrics for Kafka or Kafka Connect when multiple listeners are set in Kafka.

SMM cannot identify multiple listeners and still points to bootstrap server using the default broker port (9093 for SASL_SSL). You would have to override bootstrap server URL (hostname:port as set in the listeners for broker). Add the bootstrap server details in SMM safety valve in the following path:

Cloudera Manager SMM Configuration Streams Messaging Manager Rest Admin Server Advanced Configuration Snippet (Safety Valve) for streams-messaging-manager.yaml Add the following value for bootstrap servers Save Changes Restart SMM .

```
streams.messaging.manager.kafka.bootstrap.servers=<comma-separat
ed list of brokers>
```

OPSAPS-59597: SMM UI logs are not supported by Cloudera Manager

Cloudera Manager does not support the log type used by SMM UI.

View the SMM UI logs on the host.

Limitations

CDPD-36422: 1MB flow.snapshot freezes Safari

While importing large connector configurations, flow.snapshots reduces the usability of the Streams Messaging Manager when using Safari browser.

Use a different browser (Chrome/Firefox/Edge).

Known Issues in Streams Replication Manager

Learn about the known issues in Streams Replication Manager, the impact or changes to the functionality, and the workaround.

Known Issues

CDPD-22089: SRM does not sync re-created source topics until the offsets have caught up with target topic

Messages written to topics that were deleted and re-created are not replicated until the source topic reaches the same offset as the target topic. For example, if at the time of deletion and re-creation there are a 100 messages on the source and target clusters, new messages will only get replicated once the re-created source topic has 100 messages. This leads to messages being lost.

None

CDPD-11079: Blacklisted topics appear in the list of replicated topics

If a topic was originally replicated but was later disallowed (blacklisted), it will still appear as a replicated topic under the /remote-topics REST API endpoint. As a result, if a call is made to this endpoint, the disallowed topic will be included in the response. Additionally, the disallowed topic will also be visible in the SMM UI. However, it's Partitions and Consumer Groups will be 0, its Throughput, Replication Latency and Checkpoint Latency will show N/A.

None

CDPD-30275: SRM may automatically re-create deleted topics on target clusters

If auto.create.topics.enable is enabled, deleted topics might get automatically re-created on target clusters. This is a timing issue. It only occurs if remote topics are deleted while the replication of the topic is still ongoing.

1. Remove the topic from the topic allowlist with srm-control. For example:

```
srm-control topics --source [SOURCE_CLUSTER] --target [TARGE
T_CLUSTER] --remove [TOPIC1]
```

2. Wait until SRM is no longer replicating the topic.
3. Delete the remote topic in the target cluster.

OPSAPS-67772: SRM Service metrics processing fails when the noexec option is enabled for /tmp

The SRM Service role uses /tmp to extract RocksDB .so files, which are required for metrics processing to function. If the noexec option is enabled for the /tmp directory, the SRM Service role is not able load the required RocksDB files. This results in metrics processing failing.

1. In Cloudera Manager, select the SRM service and go to Configuration.

2. Add the following to SRM Service Environment Advanced Configuration Snippet (Safety Valve). Do this for all SRM Service role instances.

```
ROCKSDB_SHAREDLIB_DIR=[***PATH***]
```

Replace *[***PATH***]* with a directory that is not /tmp.

OPSAPS-67738: SRM Service role's Remote Querying feature does not work when the noexec option is enabled for /tmp

The SRM Service role puts the Netty native libraries into the /tmp directory. As a result, If the noex ec option is enabled for the /tmp directory, the Remote Querying feature will fail to function.

1. In Cloudera Manager, select the SRM service and go to Configuration.
2. Add the following to SRM_JVM_PERF_OPTS.

```
-Dio.netty.native.workdir=[***PATH***]
```

Replace *[***PATH***]* with a directory that is not /tmp.

OPSAPS-67742: The SRM Service role fails to start if properties are added to Additional Configs For Streams Application Running Inside SRM Service

Configuring the SRM Service role's internal Kafka Streams application is not possible. If you add any properties to Using the Additional Configs For Streams Application Running Inside SRM Service, the SRM Service role fails to start. If you are affected by this issue, an exception similar to the following will be present in the SRM Service role's stderr.log:

```
o.dropwizard.configuration.ConfigurationParsingException: /var/r
un/cloudera-scm-agent/process/132-streams_replication_manager-ST
REAMS_REPLICATION_MANAGER_SERVICE/srm-service.yaml has an error:
 * Malformed YAML at line: 66, column: 49; mapping values are not
 allowed here in 'reader', line 65, column 48:
.
.
```

None

CDPD-60426: Configuration changes are lost following a rolling restart of the service

In certain cases, SRM might fail to apply configuration updates if the service is restarted with a rolling restart. In a case like this, configuration changes are ignored without any warning or indication. This issue also affects rolling upgrades.

When restarting the service, use Actions Restart instead of Actions Rolling Restart after making configuration changes. When upgrading a cluster, ensure that SRM is not restarted with a rolling restart.

Limitations

SRM cannot replicate Ranger authorization policies to or from Kafka clusters

Due to a limitation in the Kafka-Ranger plugin, SRM cannot replicate Ranger policies to or from clusters that are configured to use Ranger for authorization. If you are using SRM to replicate data to or from a cluster that uses Ranger, disable authorization policy synchronization in SRM. This can be achieved by clearing the Sync Topic Acls Enabled (sync.topic.acls.enabled) checkbox.

Known Issues in MapReduce, Apache Hadoop YARN, and YARN Queue Manager

Learn about the known issues in Mapreduce, YARN and YARN Queue Manager, the impact or changes to the functionality, and the workaround.

Known Issues

CDPD-46685 Nodemanager logs are filled with logs similar to: 2022-11-28 03:42:39,587 WARN org.apache.hadoop.ipc.Client: Address change detected. Old: deh-34631355-niv-master1.e2e-797.dze1-y40r.int.cldr.work/10.114.128.84:8031 New: deh-34631355-niv-master1.e2e-797.dze1-y40r.int.cldr.work/10.114.128.63:8031 2022-11-28 03:43:01,425 WARN org.apache.hadoop.ipc.Client: Address change detected. Old: deh-34631355-niv-master0.e2e-797.dze1-y40r.int.cldr.work/10.114.128.79:8031 New: deh-34631355-niv-master0.e2e-797.dze1-y40r.int.cldr.work/10.114.128.65:8031.

Restart all YARN NodeManagers, they should come up without issues and Cloudera Manager should recognize them as healthy nodes once the status of them is refreshed upon restart.

YARN cannot start if Kerberos principal name is changed

If the Kerberos principal name is changed in Cloudera Manager after launch, YARN will not be able to start. In such case the keytabs can be correctly generated but YARN cannot access ZooKeeper with the new Kerberos principal name and old ACLs.

There are two possible workarounds:

- Delete the znode and restart the YARN service.
- Use the reset ZK ACLs command. This also sets the znodes below /rmstore/ZKRMStateRoot to world:anyone:cdrwa which is less secure.

Third party applications do not launch if MapReduce framework path is not included in the client configuration

MapReduce application framework is loaded from HDFS instead of being present on the NodeManagers. By default the mapreduce.application.framework.path property is set to the appropriate value, but third party applications with their own configurations will not launch.

Set the mapreduce.application.framework.path property to the appropriate configuration for third party applications.

JobHistory URL mismatch after server relocation

After moving the JobHistory Server to a new host, the URLs listed for the JobHistory Server on the ResourceManager web UI still point to the old JobHistory Server. This affects existing jobs only. New jobs started after the move are not affected.

For any existing jobs that have the incorrect JobHistory Server URL, there is no option other than to allow the jobs to roll off the history over time. For new jobs, make sure that all clients have the updated mapred-site.xml that references the correct JobHistory Server.

CDH-6808: Routable IP address required by ResourceManager

ResourceManager requires routable host:port addresses for yarn.resourcemanager.scheduler.address, and does not support using the wildcard 0.0.0.0 address.

Set the address, in the form host:port, either in the client-side configuration, or on the command line when you submit the job.

CDH-49165: History link in ResourceManager web UI broken for killed Spark applications

When a Spark application is killed, the history link in the ResourceManager web UI does not work.

To view the history for a killed Spark application, see the Spark HistoryServer web UI instead.

CDPD-2936: Application logs are not accessible in WebUI2 or Cloudera Manager

Running Containers Logs from NodeManager local directory cannot be accessed either in Cloudera Manager or in WebUI2 due to log aggregation.

Use the YARN log CLI to access application logs. For example:

```
yarn logs -applicationId <APPLICATION ID>
```

Apache Issue: YARN-9725

COMPX-1445: Queue Manager operations are failing when Queue Manager is installed separately from YARN

If Queue Manager is not selected during YARN installation, Queue Manager operation are failing. Queue Manager says 0 queues are configured and several failures are present. That is because ZooKeeper configuration store is not enabled.

1. In Cloudera Manager, select the YARN service.
2. Click the Configuration tab.
3. Find the Queue Manager Service property.
4. Select the Queue Manager service that the YARN service instance depends on.
5. Click Save Changes.
6. Restart all services that are marked stale in Cloudera Manager.

COMPX-3329: Autorestart is not enabled for Queue Manager in Data Hub

In a Data Hub cluster, Queue Manager is installed with autorestart disabled. Hence, if Queue Manager goes down, it will not restart automatically.

If Queue Manager goes down in a Data Hub cluster, you must go to the Cloudera Manager Dashboard and restart the Queue Manager service.

COMPX-5817: Queue Manager UI will not be able to present a view of pre-upgrade queue structure. CM Store is not supported and therefore Yarn will not have any of the pre-upgrade queue structure preserved.

When a Data Hub cluster is deleted, all saved configurations are also deleted. All YARN configurations are saved in CM Store and this is yet to be supported in Data Hub and Cloudera Manager. Hence, the YARN queue structure also will be lost when a Data Hub cluster is deleted or upgraded or restored.

Technical Service Bulletins

TSB 2023-641: InvalidClassException while editing queue configurations in YARN Queue Manager UI

Under situations described below, a user may encounter the following error message while editing queue configurations in the Apache Hadoop YARN (YARN) Queue Manager UI:

Failed to update queue configuration

Modify queue operation failed. Queue configuration in Cloudera Manager is inconsistent with YARN. Try restarting YARN.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: TSB 2022-641: InvalidClassException while editing queue configurations in YARN Queue Manager UI.

Unsupported Features

The following YARN features are currently not supported in Cloudera Data Platform:

- Application Timeline Server (ATSv2 and ATSv1)
- Container Resizing
- Distributed or Centralized Allocation of Opportunistic Containers
- Distributed Scheduling

- Docker on YARN (DockerContainerExecutor) on Data Hub clusters
- Fair Scheduler
- GPU support for Docker
- Hadoop Pipes
- Native Services
- Pluggable Scheduler Configuration
- Queue Priority Support
- Reservation REST APIs
- Resource Estimator Service
- Resource Profiles
- (non-Zookeeper) ResourceManager State Store
- Rolling Log Aggregation
- Shared Cache
- YARN Federation
- Moving jobs between queues

Known Issues in Apache Zeppelin

Learn about the known issues in Zeppelin, the impact or changes to the functionality, and the workaround.

TSB 2024-650: Arbitrary file deletion vulnerability in Apache Zeppelin

The improper Input Validation vulnerability in Apache Zeppelin allows an attacker to delete arbitrary files. Using a successful cross-site scripting attack by accessing the logs through API:

```
/api/interpreter/setting/..%2Flogs
```

The logs folder can be deleted from the directory where the current project is located. If the API is changed to /api/interpreter/setting/..%2F..%2Fzeppelin the following setting, the entire Zeppelin application directory can be deleted. The Zeppelin application directory contains every configuration file, Zeppelin main program files, and so on, which are crucial for the proper operations of Zeppelin.

Upstream JIRA: ZEPPELIN-5624

For the latest update on this issue see the corresponding Knowledge article: TSB 2024-650: TSB title

CDPD-3090: Due to a configuration typo, functionality involving notebook repositories does not work

Due to a missing closing brace, access to the notebook repositories API is blocked by default.

From the CDP Management Console, go to Cloudera Manager for the cluster running Zeppelin. On the Zeppelin configuration page (Zeppelin serviceConfiguration), enter shiro urls in the Search field, and then add the missing closing brace to the notebook-repositories URL, as follows:

```
/api/notebook-repositories/** = authc, roles[{{zeppelin_admin_gr
oup}}]
```

Click Save Changes, and restart the Zeppelin service.

CDPD-2406: Logout button does not work

Clicking the Logout button in the Zeppelin UI logs you out, but then immediately logs you back in using SSO.

Close the browser.

Known Issues in Apache ZooKeeper

Learn about the known issues in Zookeeper, the impact or changes to the functionality, and the workaround.

Zookeeper-client does not use ZooKeeper TLS/SSL automatically

The command-line tool 'zookeeper-client' is installed to all Cloudera Nodes and it can be used to start the default Java command line ZooKeeper client. However even when ZooKeeper TLS/SSL is enabled, the zookeeper-client command connects to localhost:2181, without using TLS/SSL.

Manually configure the 2182 port, when zookeeper-client connects to a ZooKeeper cluster.The following is an example of connecting to a specific three-node ZooKeeper cluster using TLS/SSL:

```
CLIENT_JVMFLAGS="-Dzookeeper.clientCnxnSocket=org.apache.zoo
keeper.ClientCnxnSocketNetty -Dzookeeper.ssl.keyStore.locati
on=<PATH TO YOUR CONFIGURED KEYSTORE> -Dzookeeper.ssl.keyStor
e.password=<THE PASSWORD YOU CONFIGURED FOR THE KEYSTORE>  -
Dzookeeper.ssl.trustStore.location=<PATH TO YOUR CONFIGURED
 TRUSTSTORE> -Dzookeeper.ssl.trustStore.password=<THE PASSWORD
 YOU CONFIGURED FOR THE TRUSTSTORE> -Dzookeeper.client.secu
re=true" zookeeper-client -server <YOUR.ZOOKEEPER.SERVER-1>:218
2,<YOUR.ZOOKEEPER.SERVER-2>:2182,<YOUR.ZOOKEEPER.SERVER-3>:2182
```

Behavioral Changes In Cloudera Runtime 7.2.16

You can review the changes in certain features or functionalities of components that have resulted in a change in behavior from the previously released version to this version of Cloudera Runtime 7.2.16.

Behavioral Changes in Apache Hive

Learn about the change in certain functionality of Hive that has resulted in a change in behavior from the previously released version to this version of Cloudera Runtime.

Summary:

CDP runtine 7.2.16 inroduces a new configuration property that affects how Ranger vailidates Hive url policies. Specifically, this new property allows you to restore the 7.2.15 behavior when Ranger validates Hive URL policies.

Previous behavior:

In 7.2.15, Ranger used a relative path to validate Hive url policies.

New behavior:

To configure the type of path that Ranger uses to validate Hive url policies:

1. Go to Cloudera Manager Hive_on_Tez Configuration .
2. In Search, type hive_service.
3. In Hive Service Advanced Configuration Snippet (Safety-Valve) for hive-site.xml, click +.
4. Add the file path configuration property.
   a. In Name, type hive.ranger.use.fully.qualified.url
   b. In Value, type true.
5. Click Save Changes (CRTL+S).

When you set hive.ranger.use.fully.qualified.url to true, a fully qualified path will be used to validate against ranger url policies. When you set hive.ranger.use.fully.qualified.url to false, relative path is used. You cannot modify hive.ranger.use.fully.qualified.url at runtime.

For more information about creating Hive URL policies, see Create a Hive authorizer URL policy.

Behavioral Changes in Apache Kafka

Learn about the change in certain functionality of Kafka that has resulted in a change in behavior from the previously released version to this version of Cloudera Runtime.

Summary:

The FileStream example connectors (FileStreamSourceConnector and FileStreamsSinkConnector) are no longer available for deployment by default using the Kafka Connect Rest API, SMM UI, or SMM REST API. The JAR file of the connector is still shipped with Cloudera Runtime, but the connectors must be installed before they can be deployed. For more information on how to install Kafka Connect connectors, see Installing Connectors

Previous behavior:

TheFileStream example connectors were available by default for deployment. In SMM, the connectors were selectable on the Connect ClusterConnector Setup page by default.

New behavior:

TheFileStream example connectors must be installed before they can be deployed. In SMM, the connectors are no longer selectable on the Connect ClusterConnector Setup page by default.

Summary:

Topic auto creation for the internal consumers of the Kafka Connect workers is turned off by default.

Previous behavior:

Topic auto creation for the internal consumers of the Kafka Connect workers was turned on by default.

New behavior:

Topic auto creation for the internal consumers of the Kafka Connect workers is turned off by default.

Behavioral Changes in Schema Registry

Learn about the change in certain functionality of Schema Registry that has resulted in a change in behavior from the previously released version to this version of Cloudera Runtime.

Summary:

Schema compatibility check now supports the evolution of Avro enums.

Previous behavior:

Schema compatibility check fails when a writer schema that has enum values is not present in the reader schema (some previous version), regardless if the reader has a default value.

New behavior:

A schema that does not have all enum values but has a default value set, remains compatible with a newer version that has an extended set of enum values.

Behavioral Changes in Streams Messaging Manager

Learn about the change in certain functionality of Streams Messaging Manager that has resulted in a change in behavior from the previously released version to this version of Cloudera Runtime.

Summary:

The FileStream example connectors (FileStreamSourceConnector and FileStreamsSinkConnector) are no longer available for deployment by default using the Kafka Connect Rest API, SMM UI, or SMM REST API. The JAR file of the connector is still shipped with Cloudera Runtime, but the

connectors must be installed before they can be deployed. For more information on how to install Kafka Connect connectors, see Installing Connectors

Previous behavior:

TheFileStream example connectors were available by default for deployment. In SMM, the connectors were selectable on the Connect ClusterConnector Setup page by default.

New behavior:

TheFileStream example connectors must be installed before they can be deployed. In SMM, the connectors are no longer selectable on the Connect ClusterConnector Setup page by default.

Behavioral Changes in Streams Replication Manager

Learn about the change in certain functionality of Streams Replication Manager that has resulted in a change in behavior from the previously released version to this version of Cloudera Runtime.

Summary:

The default values for Metrics Topics Creation Timeout for Driver and Metrics Topics Creation Timeout for Service are increased to 20 seconds

Previous behavior:

The default values for Metrics Topics Creation Timeout for Driver and Metrics Topics Creation Timeout for Service was 5 seconds.

New behavior:

The default values for Metrics Topics Creation Timeout for Driver and Metrics Topics Creation Timeout for Service is 20 seconds.

Deprecation Notices In Cloudera Runtime 7.2.16

Certain features and functionalities have been removed or deprecated in Cloudera Runtime 7.2.16. You must review these items to understand whether you must modify your existing configuration. You can also learn about the features that will be removed or deprecated in the future release to plan for the required changes.

Terminology

Items in this section are designated as follows:

Deprecated

Technology that Cloudera is removing in a future CDP release. Marking an item as deprecated gives you time to plan for removal in a future CDP release.

Moving

Technology that Cloudera is moving from a future CDP release and is making available through an alternative Cloudera offering or subscription. Marking an item as moving gives you time to plan for removal in a future CDP release and plan for the alternative Cloudera offering or subscription for the technology.

Removed

Technology that Cloudera has removed from CDP and is no longer available or supported as of this release. Take note of technology marked as removed since it can potentially affect your upgrade plans.

Removed Components and Product Capabilities

No components are deprecated or removed in this Cloudera Runtime release.

Please contact Cloudera Support or your Cloudera Account Team if you have any questions.

Deprecation Notices for Apache Kafka

Certain features and functionality in Apache Kafka are deprecated or removed in Cloudera Runtime 7.2.16. You must review these changes along with the information about the features in Kafka that will be removed or deprecated in a future release.

Important: The following list of deprecated and removed items is not exhaustive and only contains items that have a direct and immediate effect on Kafka in CDP. For a full list of deprecation and/or removals in the version Apache Kafka shipped with Runtime, review the *Notable Changes* as well as the *Release Notes* on https://kafka.apache.org/.

Removed

kafka-preferred-replica-election

The kafka-preferred-replica-election.sh command line tool is removed. Its alternative in CDP, kafka-preferred-replica-election, is also removed. Use kafka-leader-election instead.

--zookeeper

The --zookeeper option is removed from the kafka-topics and kafka-reassign-partitions command line tools. Use the --bootstrap-server option instead.

Removd properties

The following Kafka service properties are removed:

- Default Consumer Quota (quota.consumer.default)
- Default Producer Quota (quota.producer.default)
- Advertised Host (advertised.port)
- Advertised Port (advertised.host.name)
- Kafka Connect Prometheus Metrics Port (connect.prometheus.metrics.port)

Deprecated

MirrorMaker (MM1)

MirrorMaker is deprecated. Cloudera recommends that you use Streams Replication Manager (SRM) instead.

--zookeeper

The --zookeeper option is only supported for the kafka-configs tool and should be only used when updating SCRAM Credential configurations. The --zookeeper option is either deprecated in or removed from other Kafka command line tools. Cloudera recommends that you use the --bootstrap-server option instead.

Fixed Common Vulnerabilities and Exposures 7.2.16

Common Vulnerabilities and Exposures (CVE) that is fixed in this release.

- CVE-2022-40664
- CVE-2022-42889
- CVE-2021-27568
- CVE-2021-22144
- CVE-2021-22135
- CVE-2021-22137
- CVE-2021-36373

- CVE-2021-36374
- CVE-2020-11987
- CVE-2020-27568
- CVE-2019-5427
- CVE-2020-13955
- CVE-2018-11771
- CVE-2021-41269
- CVE-2021-40690
- CVE-2021-26919
- CVE-2021-26920
- CVE-2021-36749
- CVE-2021-44791
- CVE-2022-28889
- CVE-2018-1000840
- CVE-2020-17521
- CVE-2015-6584
- CVE-2017-7536
- CVE-2018-18928
- CVE-2020-10531
- CVE-2020-21913
- CVE-2021-28165
- CVE-2021-28164
- CVE-2021-28163
- CVE-2021-37714
- CVE-2022-23596
- CVE-2020-29582
- CVE-2022-24329
- CVE-2021-20218
- CVE-2018-1320
- CVE-2019-0205
- CVE-2019-0210
- CVE-2018-11798
- CVE-2016-5397
- CVE-2017-12629
- CVE-2022-24613
- CVE-2022-24614
- CVE-2021-20328
- CVE-2016-2402
- CVE-2021-27807
- CVE-2021-27906
- CVE-2021-31811
- CVE-2021-31812
- CVE-2022-26336
- CVE-2022-21724
- CVE-2022-26520
- CVE-2022-34169
- CVE-2022-32532
- CVE-2022-22965
- CVE-2022-22950
- CVE-2022-22968

- CVE-2022-22970
- CVE-2022-22971
- CVE-2021-22060
- CVE-2021-22096
- CVE-2021-22118
- CVE-2022-22112
- CVE-2021-22118
- CVE-2016-22118
- CVE-2016-9879
- CVE-2019-11272
- CVE-2021-9879
- CVE-2022-22976
- CVE-2022-25169
- CVE-2022-30126
- CVE-2022-30973
- CVE-2022-33879
- CVE-2022-25762
- CVE-2022-29885
- CVE-2022-23181
- CVE-2020-13936
- CVE-2022-23437
- CVE-2020-11988