Cloudera Runtime 7.2.14

# Release Notes

**Date published: 2022-02-24**
**Date modified:**

**https://docs.cloudera.com/**

# Legal Notice

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

# Contents

## Behavioral Changes In Cloudera Runtime 7.2.14......87

## Deprecation Notices In Cloudera Runtime 7.2.14......89

# Overview

You can review the Release Notes of Cloudera Runtime 7.2.14 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 7.2.14 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 | 0.11.1.7.2.14.0-149 |
| Apache Atlas | 2.1.0.7.2.14.0-149 |
| Apache Calcite | 1.21.0.7.2.14.0-149 |
| Apache Avro | 1.8.2.7.2.14.0-149 |
| Apache Hadoop (Includes YARN and HDFS) | 3.1.1.7.2.14.0-149 |
| Apache HBase | 2.4.6.7.2.14.0-149 |
| Apache Hive | 3.1.3000.7.2.14.0-149 |
| Apache Impala | 4.0.0.7.2.14.0-149 |
| Apache Kafka | 2.8.1.7.2.14.0-149 |
| Apache Knox | 1.3.0.7.2.14.0-149 |
| Apache Kudu | 1.15.0.7.2.14.0-149 |
| Apache Livy | 0.6.0.7.2.14.0-149 |
| Apache MapReduce | 3.1.1.7.2.14.0-149 |
| Apache NiFi | 1.15.0.7.2.14.0-149 |
| Apache NiFi Registry | 1.15.2.7.2.14.0-149 |
| Apache Oozie | 5.1.0.7.2.14.0-149 |
| Apache ORC | 1.5.1.7.2.14.0-149 |
| Apache Parquet | 1.10.99.7.2.14.0-149 |
| Apache Phoenix | 5.1.1.7.2.14.0-149 |
| Apache Ranger | 2.1.0.7.2.14.0-149 |
| Apache Solr | 8.4.1.7.2.14.0-149 |
| Apache Spark | 2.4.8.7.2.14.0-149 |
| Apache Sqoop | 1.4.7.7.2.14.0-149 |
| Apache Tez | 0.9.1.7.2.14.0-149 |
| Apache Zeppelin | 0.8.2.7.2.14.0-149 |
| Apache ZooKeeper | 3.5.5.7.2.14.0-149 |

Other Components

| Component | Version |
|---|---|
| Cruise Control | 2.5.66.7.1.7.0-551 |
| Data Analytics Studio | 1.4.2.7.2.14.0-149 |
| GCS Connector | 2.1.2.7.2.14.0-149 |
| HBase Indexer | 1.5.0.7.2.14.0-149 |
| Hue | 4.5.0.7.2.14.0-149 |
| Search | 1.0.0.7.2.14.0-149 |
| Schema Registry | 0.10.0.7.2.14.0-149 |
| Streams Messaging Manager | 2.2.0.7.2.14.0-149 |
| Streams Replication Manager | 1.1.0.7.2.14.0-149 |

Connectors and Encryption Components

| Component | Version |
|---|---|
| HBase connectors | 1.0.0.7.2.14.0-149 |
| Hive Meta Store (HMS) | 1.0.0.7.2.14.0-149 |
| Hive on Tez | 1.0.0.7.2.14.0-149 |
| Hive Warehouse Connector | 1.0.0.7.2.14.0-149 |
| Spark Atlas Connector | 0.1.0.7.2.14.0-149 |
| Spark Schema Registry | 1.1.0.7.2.14.0-149 |

# Using the Cloudera Runtime Maven repository

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 7.2.14

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 | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-aws-s3-bridge | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-azure-adls-bridge | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-classification-updater | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-client-common | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-client-v1 | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-client-v2 | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-common | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-distro | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-docs | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-graphdb-api | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-graphdb-common | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-graphdb-janus | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-index-repair-tool | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-intg | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-janusgraph-hbase2 | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-notification | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-plugin-classloader | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-repository | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-server-api | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | atlas-testtools | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | hbase-bridge | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | hbase-bridge-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | hbase-testing-util | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | hdfs-model | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | hive-bridge | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | hive-bridge-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | impala-bridge | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | impala-bridge-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | impala-hook-api | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | kafka-bridge | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | kafka-bridge-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | navigator-to-atlas | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | sample-app | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | sqoop-bridge | 2.1.0.7.2.14.0-149 |
| | org.apache.atlas | sqoop-bridge-shim | 2.1.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| Apache Avro | org.apache.avro | avro | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-compiler | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-ipc | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-mapred | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-maven-plugin | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-protobuf | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-service-archetype | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-thrift | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | avro-tools | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | trevni-avro | 1.8.2.7.2.14.0-149 |
| | org.apache.avro | trevni-core | 1.8.2.7.2.14.0-149 |
| Apache Calcite | org.apache.calcite | calcite-babel | 1.21.0.7.2.14.0-149 |
| | org.apache.calcite | calcite-core | 1.21.0.7.2.14.0-149 |
| | org.apache.calcite | calcite-druid | 1.21.0.7.2.14.0-149 |
| | org.apache.calcite | calcite-kafka | 1.21.0.7.2.14.0-149 |
| | org.apache.calcite | calcite-linq4j | 1.21.0.7.2.14.0-149 |
| | org.apache.calcite | calcite-server | 1.21.0.7.2.14.0-149 |
| | org.apache.calcite | calcite-ubenchmark | 1.21.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica | 1.16.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica-core | 1.16.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica-metrics | 1.16.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica-metrics-dropwizardmetrics | 1.16.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica-noop-driver | 1.16.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica-server | 1.16.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica-standalone-server | 1.16.0.7.2.14.0-149 |
| | org.apache.calcite.avatica | avatica-tck | 1.16.0.7.2.14.0-149 |
| Apache Druid | org.apache.druid | druid-aws-common | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-benchmarks | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-console | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-core | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-gcp-common | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-hll | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-indexing-hadoop | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-indexing-service | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-integration-tests | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-processing | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-server | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-services | 0.17.1.7.2.14.0-149 |
| | org.apache.druid | druid-sql | 0.17.1.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.druid | extendedset | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-avro-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-basic-security | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-bloom-filter | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-datasketches | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-ec2-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-google-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-hdfs-storage | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-histogram | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-kafka-extraction-namespace | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-kafka-indexing-service | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-kerberos | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-kinesis-indexing-service | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-lookups-cached-global | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-lookups-cached-single | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-orc-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-parquet-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-protobuf-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-s3-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | druid-stats | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | mysql-metadata-storage | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | postgresql-metadata-storage | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions | simple-client-sslcontext | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | ambari-metrics-emitter | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | dropwizard-emitter | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-azure-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-cassandra-storage | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-cloudfiles-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-distinctcount | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-influx-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-influxdb-emitter | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-momentsketch | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-moving-average-query | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-opentsdb-emitter | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-redis-cache | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-tdigestsketch | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-thrift-extensions | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-time-min-max | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | druid-virtual-columns | 0.17.1.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.druid.extensions.contrib | graphite-emitter | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | kafka-emitter | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | materialized-view-maintenance | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | materialized-view-selection | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | sqlserver-metadata-storage | 0.17.1.7.2.14.0-149 |
| | org.apache.druid.extensions.contrib | statsd-emitter | 0.17.1.7.2.14.0-149 |
| GCS Connector | com.google.cloud.bigdataoss | bigquery-connector | 2.1.2.7.2.14.0-149 |
| | com.google.cloud.bigdataoss | gcs-connector | 2.1.2.7.2.14.0-149 |
| | com.google.cloud.bigdataoss | gcsio | 2.1.2.7.2.14.0-149 |
| | com.google.cloud.bigdataoss | util | 2.1.2.7.2.14.0-149 |
| | com.google.cloud.bigdataoss | util-hadoop | 2.1.2.7.2.14.0-149 |
| Apache Hadoop | org.apache.hadoop | hadoop-aliyun | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-annotations | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-archive-logs | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-archives | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-assemblies | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-auth | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-aws | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-azure | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-azure-datalake | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-build-tools | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-client | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-client-api | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-client-integration-tests | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-client-minicluster | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-client-runtime | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-cloud-storage | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-common | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-datajoin | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-distcp | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-extras | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-fs2img | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-gridmix | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-client | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-common | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-config | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-container-service | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-docs | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-hadoop-dependency-client | 1.1.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.hadoop | hadoop-hdds-hadoop-dependency-server | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-hadoop-dependency-test | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-interface-admin | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-interface-client | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-interface-server | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-server-framework | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-server-scm | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-test-utils | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdds-tools | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdfs | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdfs-client | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdfs-httpfs | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdfs-native-client | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdfs-nfs | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-hdfs-rbf | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-kafka | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-kms | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-app | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-common | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-core | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-hs | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-hs-plugins | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-jobclient | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-nativetask | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-shuffle | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-client-uploader | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-mapreduce-examples | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-maven-plugins | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-minicluster | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-minikdc | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-nfs | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-openstack | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-client | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-common | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-csi | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-datanode | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-dist | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-filesystem | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-filesystem-common | 1.1.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.hadoop | hadoop-ozone-filesystem-hadoop2 | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-filesystem-hadoop3 | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-filesystem-shaded | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-insight | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-integration-test | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-interface-client | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-interface-storage | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-network-tests | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-ozone-manager | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-recon | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-reconcodegen | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-s3gateway | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-ozone-tools | 1.1.0.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-resourceestimator | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-rumen | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-sls | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-streaming | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-tools-dist | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-api | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-applications-distributedshell | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-applications-unmanaged-am-launcher | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-client | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-common | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-registry | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-applicationhistoryservice | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-common | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-nodemanager | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-resourcemanager | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-router | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-sharedcachemanager | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-tests | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-timeline-pluginstorage | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-timelineservice | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-timelineservice-hbase-client | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-timelineservice-hbase-common | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-timelineservice-hbase-server-2 | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-timelineservice-hbase-tests | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-server-web-proxy | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | hadoop-yarn-services-api | 3.1.1.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.hadoop | hadoop-yarn-services-core | 3.1.1.7.2.14.0-149 |
| | org.apache.hadoop | mini-chaos-tests | 1.1.0.7.2.14.0-149 |
| Apache HBase | org.apache.hbase | filesystem | hadoop3-3-testutils |
| | org.apache.hbase | hbase-annotations | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-asyncfs | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-checkstyle | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-client | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-client-project | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-common | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-endpoint | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-examples | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-external-blockcache | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-hadoop-compat | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-hadoop2-compat | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-hbtop | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-http | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-it | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-logging | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-mapreduce | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-metrics | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-metrics-api | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-procedure | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-protocol | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-protocol-shaded | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-replication | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-resource-bundle | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-rest | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-rsgroup | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-server | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-shaded-client | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-shaded-client-byo-hadoop | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-shaded-client-project | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-shaded-mapreduce | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-shaded-testing-util | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-shaded-testing-util-tester | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-shell | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-testing-util | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-thrift | 2.4.6.7.2.14.0-149 |
| | org.apache.hbase | hbase-zookeeper | 2.4.6.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.hbase.connectors.kafka | hbase-kafka-model | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.kafka | hbase-kafka-proxy | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark-it | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark-protocol | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark-protocol-shaded | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark3 | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark3-it | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark3-protocol | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.connectors.spark | hbase-spark3-protocol-shaded | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.filesystem | hadoop-testutils | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.filesystem | hbase-fs-impl | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.filesystem | hbase-oss | 1.0.0.7.2.14.0-149 |
| | org.apache.hbase.thirdparty | hbase-noop-htrace | 3.5.0.7.2.14.0-149 |
| | org.apache.hbase.thirdparty | hbase-shaded-gson | 3.5.0.7.2.14.0-149 |
| | org.apache.hbase.thirdparty | hbase-shaded-jersey | 3.5.0.7.2.14.0-149 |
| | org.apache.hbase.thirdparty | hbase-shaded-jetty | 3.5.0.7.2.14.0-149 |
| | org.apache.hbase.thirdparty | hbase-shaded-miscellaneous | 3.5.0.7.2.14.0-149 |
| | org.apache.hbase.thirdparty | hbase-shaded-netty | 3.5.0.7.2.14.0-149 |
| | org.apache.hbase.thirdparty | hbase-shaded-protobuf | 3.5.0.7.2.14.0-149 |
| Apache Hive | org.apache.hive | catalogd-unit | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-beeline | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-blobstore | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-classification | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-cli | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-common | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-contrib | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-druid-handler | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-exec | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-hbase-handler | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-hcatalog-it-unit | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-hplsql | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-impala | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-custom-serde | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-druid | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-impala | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-minikdc | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-qfile | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-qfile-kudu | 3.1.3000.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.hive | hive-it-test-serde | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-unit | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-unit-hadoop2 | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-it-util | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-jdbc | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-jdbc-handler | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-jmh | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-kryo-registrator | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-kudu-handler | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-llap-client | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-llap-common | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-llap-ext-client | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-llap-server | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-llap-tez | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-metastore | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-parser | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-pre-upgrade | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-serde | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-service | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-service-rpc | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-shims | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-spark-client | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-standalone-metastore | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-storage-api | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-streaming | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-testutils | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-udf | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | hive-vector-code-gen | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive | kafka-handler | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hcatalog | hive-hcatalog-core | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hcatalog | hive-hcatalog-pig-adapter | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hcatalog | hive-hcatalog-server-extensions | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hcatalog | hive-hcatalog-streaming | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hcatalog | hive-webhcat | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hcatalog | hive-webhcat-java-client | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hive-it-custom-udfs | udf-classloader-udf1 | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hive-it-custom-udfs | udf-classloader-udf2 | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.hive-it-custom-udfs | udf-classloader-util | 3.1.3000.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.hive.hive-it-custom-udfs | udf-vectorized-badexample | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.shims | hive-shims-0.23 | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.shims | hive-shims-common | 3.1.3000.7.2.14.0-149 |
| | org.apache.hive.shims | hive-shims-scheduler | 3.1.3000.7.2.14.0-149 |
| Apache Hive Warehouse Connector | com.hortonworks.hive | hive-warehouse-connector_2.11 | 1.0.0.7.2.14.0-149 |
| Apache Kafka | org.apache.kafka | connect | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-api | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-basic-auth-extension | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-cloudera-authorization-extension | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-file | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-json | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-mirror | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-mirror-client | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-runtime | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | connect-transforms | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | generator | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | jmh-benchmarks | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-clients | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-cloudera-metrics-reporter_2.12 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-cloudera-metrics-reporter_2.13 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-cloudera-plugins | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-examples | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-log4j-appender | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-metadata | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-raft | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-shell | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-examples | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-scala_2.12 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-scala_2.13 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-test-utils | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-0100 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-0101 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-0102 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-0110 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-10 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-11 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-20 | 2.8.1.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.kafka | kafka-streams-upgrade-system-tests-21 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-22 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-23 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-24 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-25 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-26 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-streams-upgrade-system-tests-27 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka-tools | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka_2.12 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | kafka_2.13 | 2.8.1.7.2.14.0-149 |
| | org.apache.kafka | ranger-kafka-connect-plugin | 2.8.1.7.2.14.0-149 |
| Apache Knox | org.apache.knox | gateway-adapter | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-admin-ui | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-applications | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-cloud-bindings | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-demo-ldap | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-demo-ldap-launcher | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-discovery-ambari | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-discovery-cm | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-docker | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-i18n | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-i18n-logging-log4j | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-i18n-logging-sl4j | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-performance-test | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-ha | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-identity-assertion-common | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-identity-assertion-concat | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-identity-assertion-hadoop-groups | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-identity-assertion-pseudo | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-identity-assertion-regex | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-identity-assertion-switchcase | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-jersey | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-rewrite | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-rewrite-common | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-rewrite-func-hostmap-static | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-rewrite-func-inbound-query-param | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-rewrite-func-service-registry | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-rewrite-step-encrypt-uri | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-rewrite-step-secure-query | 1.3.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.knox | gateway-provider-security-authc-anon | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-authz-acls | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-authz-composite | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-clientcert | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-hadoopauth | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-jwt | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-pac4j | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-preauth | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-shiro | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-provider-security-webappsec | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-release | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-server | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-server-launcher | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-server-xforwarded-filter | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-admin | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-as | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-definitions | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-hashicorp-vault | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-hbase | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-health | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-hive | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-idbroker | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-impala | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-jkg | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-knoxsso | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-knoxssout | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-knoxtoken | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-livy | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-metadata | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-nifi | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-nifi-registry | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-remoteconfig | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-rm | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-session | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-storm | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-test | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-tgs | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-vault | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-service-webhdfs | 1.3.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.knox | gateway-shell | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-shell-launcher | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-shell-release | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-shell-samples | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-spi | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-test | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-test-idbroker | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-test-release-utils | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-test-utils | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-topology-hadoop-xml | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-topology-simple | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-util-common | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-util-configinjector | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-util-launcher | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | gateway-util-urltemplate | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | hadoop-examples | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | knox-cli-launcher | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | knox-homepage-ui | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | knox-token-management-ui | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | webhdfs-kerb-test | 1.3.0.7.2.14.0-149 |
| | org.apache.knox | webhdfs-test | 1.3.0.7.2.14.0-149 |
| Apache Kudu | org.apache.kudu | kudu-backup-tools | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-backup2_2.11 | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-backup3_2.12 | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-client | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-hive | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-spark2-tools_2.11 | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-spark2_2.11 | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-spark3-tools_2.12 | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-spark3_2.12 | 1.15.0.7.2.14.0-149 |
| | org.apache.kudu | kudu-test-utils | 1.15.0.7.2.14.0-149 |
| Apache Livy | org.apache.livy | livy-api | 0.6.0.7.2.14.0-149 |
| | org.apache.livy | livy-api | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-client-common | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-client-http | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-core_2.11 | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-core_2.12 | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-examples | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-integration-test | 0.6.3000.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.livy | livy-repl_2.11 | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-repl_2.12 | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-rsc | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-scala-api_2.11 | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-scala-api_2.12 | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-server | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-test-lib | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-thriftserver | 0.6.3000.7.2.14.0-149 |
| | org.apache.livy | livy-thriftserver-session | 0.6.3000.7.2.14.0-149 |
| Apache Lucene | org.apache.lucene | lucene-analyzers-common | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-icu | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-kuromoji | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-morfologik | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-nori | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-opennlp | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-phonetic | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-smartcn | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-analyzers-stempel | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-backward-codecs | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-benchmark | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-classification | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-codecs | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-core | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-demo | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-expressions | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-facet | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-grouping | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-highlighter | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-join | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-memory | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-misc | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-monitor | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-queries | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-queryparser | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-replicator | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-sandbox | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-spatial | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-spatial-extras | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-spatial3d | 8.4.1.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.lucene | lucene-suggest | 8.4.1.7.2.14.0-149 |
| | org.apache.lucene | lucene-test-framework | 8.4.1.7.2.14.0-149 |
| Apache Oozie | org.apache.oozie | oozie-client | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-core | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-distro | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-examples | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-fluent-job-api | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-fluent-job-client | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-server | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-distcp | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-git | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-hcatalog | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-hive | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-hive2 | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-oozie | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-spark | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-sqoop | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-sharelib-streaming | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-tools | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie | oozie-zookeeper-security-tests | 5.1.0.7.2.14.0-149 |
| | org.apache.oozie.test | oozie-mini | 5.1.0.7.2.14.0-149 |
| Apache ORC | org.apache.orc | orc-core | 1.5.1.7.2.14.0-149 |
| | org.apache.orc | orc-examples | 1.5.1.7.2.14.0-149 |
| | org.apache.orc | orc-mapreduce | 1.5.1.7.2.14.0-149 |
| | org.apache.orc | orc-shims | 1.5.1.7.2.14.0-149 |
| | org.apache.orc | orc-tools | 1.5.1.7.2.14.0-149 |
| Apache Parquet | org.apache.parquet | parquet-avro | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-cascading | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-cascading3 | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-column | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-common | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-encoding | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-format-structures | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-generator | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-hadoop | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-hadoop-bundle | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-jackson | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-pig | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-pig-bundle | 1.10.99.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.parquet | parquet-protobuf | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-scala_2.10 | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-thrift | 1.10.99.7.2.14.0-149 |
| | org.apache.parquet | parquet-tools | 1.10.99.7.2.14.0-149 |
| Apache Phoenix | org.apache.phoenix | phoenix-client-embedded-hbase-2.4 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-client-hbase-2.4 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-connectors-phoenix5-compat | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-core | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-hbase-compat-2.1.6 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-hbase-compat-2.2.5 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-hbase-compat-2.3.0 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-hbase-compat-2.4.0 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-hbase-compat-2.4.1 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-pherf | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-queryserver | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-queryserver-client | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-queryserver-it | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-queryserver-load-balancer | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-queryserver-orchestrator | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-server-hbase-2.4 | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix-tracing-webapp | 5.1.1.7.2.14.0-149 |
| | org.apache.phoenix | phoenix5-hive | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix5-hive-shaded | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix5-spark | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix | phoenix5-spark-shaded | 6.0.0.7.2.14.0-149 |
| | org.apache.phoenix.thirdparty | phoenix-shaded-commons-cli | 1.1.0.7.2.14.0-149 |
| | org.apache.phoenix.thirdparty | phoenix-shaded-guava | 1.1.0.7.2.14.0-149 |
| Apache Ranger | org.apache.ranger | conditions-enrichers | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | credentialbuilder | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | embeddedwebserver | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | jisql | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ldapconfigcheck | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-adls-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-atlas-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-atlas-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-authn | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-distro | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-examples-distro | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-hbase-plugin | 2.1.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.ranger | ranger-hbase-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-hdfs-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-hdfs-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-hive-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-hive-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-intg | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kafka-connect-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kafka-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kafka-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kms | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kms-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kms-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-knox-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-knox-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kudu-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kylin-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-kylin-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-nifi-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-nifi-registry-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-ozone-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-ozone-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-plugin-classloader | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-plugins-audit | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-plugins-common | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-plugins-cred | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-plugins-installer | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-raz-adls | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-raz-hook-abfs | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-raz-hook-s3 | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-raz-intg | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-raz-processor | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-raz-s3 | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-raz-s3-lib | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-rms-common | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-rms-hive | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-rms-plugins-common | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-rms-webapp | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-s3-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-sampleapp-plugin | 2.1.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.ranger | ranger-schema-registry-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-solr-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-solr-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-sqoop-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-sqoop-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-storm-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-storm-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-tagsync | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-tools | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-util | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-yarn-plugin | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ranger-yarn-plugin-shim | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | sample-client | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | sampleapp | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | shaded-raz-hook-abfs | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | shaded-raz-hook-s3 | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | ugsync-util | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | unixauthclient | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | unixauthservice | 2.1.0.7.2.14.0-149 |
| | org.apache.ranger | unixusersync | 2.1.0.7.2.14.0-149 |
| Apache Solr | org.apache.solr | solr-analysis-extras | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-analytics | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-cell | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-clustering | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-core | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-dataimporthandler | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-dataimporthandler-extras | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-jaegertracer-configurator | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-langid | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-ltr | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-prometheus-exporter | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-security-util | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-solrj | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-test-framework | 8.4.1.7.2.14.0-149 |
| | org.apache.solr | solr-velocity | 8.4.1.7.2.14.0-149 |
| Apache Spark | org.apache.spark | spark-avro_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-avro_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-catalyst_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-catalyst_2.12 | 3.2.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.spark | spark-core_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-core_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-graphx_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-graphx_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-hadoop-cloud_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-hadoop-cloud_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-hive_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-hive_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-kubernetes_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-kubernetes_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-kvstore_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-kvstore_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-launcher_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-launcher_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-mllib-local_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-mllib-local_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-mllib_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-mllib_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-network-common_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-network-common_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-network-shuffle_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-network-shuffle_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-network-yarn_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-network-yarn_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-repl_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-repl_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-shaded-raz | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-sketch_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-sketch_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-sql-kafka-0-10_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-sql-kafka-0-10_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-sql_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-sql_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-streaming-kafka-0-10-assembly_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-streaming-kafka-0-10-assembly_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-streaming-kafka-0-10_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-streaming-kafka-0-10_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-streaming_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-streaming_2.12 | 3.2.0.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.spark | spark-tags_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-tags_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-token-provider-kafka-0-10_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-token-provider-kafka-0-10_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-unsafe_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-unsafe_2.12 | 3.2.0.7.2.14.0-149 |
| | org.apache.spark | spark-yarn_2.11 | 2.4.8.7.2.14.0-149 |
| | org.apache.spark | spark-yarn_2.12 | 3.2.0.7.2.14.0-149 |
| Apache Sqoop | org.apache.sqoop | sqoop | 1.4.7.7.2.14.0-149 |
| | org.apache.sqoop | sqoop-test | 1.4.7.7.2.14.0-149 |
| Apache Tez | org.apache.tez | hadoop-shim | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | hadoop-shim-2.8 | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-api | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-aux-services | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-common | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-dag | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-examples | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-ext-service-tests | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-history-parser | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-javadoc-tools | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-job-analyzer | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-mapreduce | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-protobuf-history-plugin | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-runtime-internals | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-runtime-library | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-tests | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-yarn-timeline-cache-plugin | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-yarn-timeline-history | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-yarn-timeline-history-with-acls | 0.9.1.7.2.14.0-149 |
| | org.apache.tez | tez-yarn-timeline-history-with-fs | 0.9.1.7.2.14.0-149 |
| Apache Zeppelin | org.apache.zeppelin | zeppelin-angular | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-display | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-interpreter | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-jdbc | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-jupyter | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-livy | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-markdown | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-server | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-shaded-raz | 0.8.2.7.2.14.0-149 |

| Project | groupId | artifactId | version |
|---|---|---|---|
| | org.apache.zeppelin | zeppelin-shell | 0.8.2.7.2.14.0-149 |
| | org.apache.zeppelin | zeppelin-zengine | 0.8.2.7.2.14.0-149 |
| Apache ZooKeeper | org.apache.zookeeper | zookeeper | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-client-c | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-contrib-loggraph | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-contrib-rest | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-contrib-zooinspector | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-docs | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-jute | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-recipes-election | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-recipes-lock | 3.5.5.7.2.14.0-149 |
| | org.apache.zookeeper | zookeeper-recipes-queue | 3.5.5.7.2.14.0-149 |

# What's New In Cloudera Runtime 7.2.14

You must be aware of the additional functionalities and improvements to features of components in Cloudera Runtime 7.2.14. Learn how the new features and improvements benefit you.

## What's New in Cruise Control

Learn about the new features of Cruise Control in Cloudera Runtime 7.2.14.

### Cruise Control 2.5.66 Rebase

Cruise Control in Cloudera Runtime is rebased from 2.0.100 to the 2.5.66 version. The main feature changes include ZooKeeper TLS/SSL support and the Cruise Control Metric Reporter support.

### ZooKeeper TLS/SSL support for Cruise Control

When TLS is enabled on the cluster, Cruise Control automatically uses the Zookeeper for secure communication.

### Cruise Control Metric Reporter support

Beside the Cloudera Manager Metrics Reporter, the Kafka based Cruise Control Metrics Reporter can also be used. The configuration needs to be set manually, and further adjustments are needed when changing the default Metrics Reporter.

## Cruise Control Rebase Summary

In CDP Public Cloud 7.2.14, Cruise Control is rebased from 2.0.100 to the 2.5.66 version. Other than the added new feature, several issues are fixed and several features are enhanced to have a better perfomance when using Cruise Control.

### Table 1: Fixed Issues

| | |
|---|---|
| PR-1184 Fix the bug in replica movement strategy chaining | PR-1291 Fix GOAL_VIOLATION detector getting stuck execution in GENERATING_PROPOSALS_FOR_EXECUTION state |

| | |
|---|---|
| PR-1209Fix NPE if task execution takes longer than executionProgressCheckIntervalMs() | PR-1381 Fix a bug that might cause invalid throttle replica list to be used |
| PR-1231 Fix reported balancedness when there are offline brokers | PR-1476 Fix miscalculated recommendation for the number of racks to drop for clusters over-provisioned wrt rack count |
| PR-1232 Fix returns for completed_with_error tasks on user_tasks endpoint when json=false | PR-1597 Fix incorrect values generated for number of replicas by topic |
| PR-1238 Fix inconsistent/bad response in monitor substate | PR-1616 Fix throttler quota removal for in-progress tasks |
| PR-1279 Fix missed broker failure detection/self-healing upon bootstrap | PR-1676 Fix EnvConfigProvider to work well if there is no pre configured env vars |

### Table 2: Version Update

| |
|---|
| PR-1233 Upgrade to Kafka 2.5.0 in development branch migrate_to_kafka_2_5 |
| PR-1311 Add support for Kafka 2.6 brokers |
| PR-1471 Add support for Kafka version 2.7 |
| PR-1612 Upgrade to Kafka 2.8 libraries |
| PR-1614 Updated to Scala 2.13 |

### Table 3: Feature Support

| | |
|---|---|
| PR-1159 Support SPNEGO and trusted proxy authentication | PR-1569 Add support to switch from ZK to Kafka Admin Client for topic config provider class |
| PR-1245, PR-1320 Bump up ZK session and connection timeout | PR-1583 Ability to configure TLS protocols and ciphers via configuration |
| PR-1510 Add support for Alerta.io notifications | PR-1661 Support connecting to ZooKeeper with TLS |
| PR-1525 Add support for (At/Under)MinISR-based throttling/ cancellation | PR-1703 Add ZK TLS support with properties file and modify broker failure detection |

### Table 4: Goal Improvements

| | |
|---|---|
| PR-1203 Add missing sanity check for goals | PR-1400 Add gap-based balance limits for TopicReplicaDistributionGoal |
| PR-1267 LeaderReplicaDistributionGoal should honor excludedTopics during leadership movement | PR-1420 Relax low resource utlization upper limit for resource distribution goals |
| PR-1306 Update min valid windows required to start self-healing with goals using resource history | PR-1429 Add a new hard goal that ensures that each alive broker has a leader replica from a configured pool of topics |
| PR-1324 Add support for goal-based operations via maintenance event | PR-1500 Drop enforcement that anomaly.detection.goals must be a subset of self.healing.goals (if non-empty) |
| PR-1345 Add a new hard goal that evenly distributes replicas over racks | PR-1514 Prevent ResourceDistributionGoal from generating provision recommendations with a negative number of brokers to remove |
| PR-1383 Ensure that topology distribution goals compute balance constraints properly | PR-1564 Add timers to track goal violation detection and fix generation for self-healing |
| PR-1385 Add support to switch from ZK to Kafka Admin Client for topic config provider class | |

### Table 5: Functonality Enhancements

| | | |
|---|---|---|
| PR-868 Provide capacity stats for a broker | PR-1302 Support stop ongoing executions with rollback | PR-1448 Reset the provision status to ensure freshness for consecutive optimizations |
| PR-1177 Make TopicReplicationFactorAnomalyFinder ignore topic with large minISR | PR-1313 Make min execution progress check interval and slow task alerting backoff configurable | PR-1456 Update slow broker detection sensitivity and reporting details |

| | | |
|---|---|---|
| PR-1180 Update min valid windows required to start self-healing with goals using resource history | PR-1316 Added option to configure metrics topic minISR | PR-1460 Add a config for admin client request timeout |
| PR-1186 Further filter detected slow broker against pre-defined flush time threshold | PR-1332 Support handling planned maintenance events submitted via a topic | PR-1463 Add sensors to report the number of slow brokers |
| PR-1190 [ccclient-1.1.0] Add force_stop parameter | PR-1334 [ccclient-1.1.1] Extend support for anomaly detectors | PR-1469 Report ongoing replica reassignments started by an external agent |
| PR-1196 Adopt admin client-based replica reassignment API for Kafka 2.4+ | PR-1341 Adopt a shared AdminClient across selected CC components | PR-1470 Provide recommendations on the estimated resource requirements |
| PR-1198 Support environment variable resolution in configs | PR-1349 Make MetricSampler more extensible | PR-1484 Add a sensor to indicate the metadata factor of the managed Kafka cluster |
| PR-1199 Add rack information to load response | PR-1357 Add check to validate that time range start time is smaller than end time | PR-1485 Add a sensor to indicate if the cluster has partitions with RF > the number of eligible racks |
| PR-1212 Make CPU capacity threshold stricter | PR-1358 Check whether a cluster is using JBOD when populate_disk_info is true | PR-1496 Stop execution before shutting down Cruise Control |
| PR-1214 Update the Balancedness Score under Unhealthy Cluster State | PR-1360 Handle Wrapped AdminClient Timeouts as Timeouts | PR-1505 Make CC inter-broker replica reassignments resilient against partitions with ISR set > replica set |
| PR-1220 Enable configurable backoff on metrics topic creation | PR-1362 Prevent MaintenanceEventTopicReader from prematurely closing the adminClient | PR-1509 Add a sensor to identify if the cluster has partitions with ISR > replicas |
| PR-1223 Ensure that requests to update replication factor cannot cause a deadlock | PR-1364 Let implementations of OptimizationOptionsGenerator be configured with an AdminClient | PR-1533 Avoid NPE due to misused rebalance_disk parameter |
| PR-1225 Allow customization of CruiseControlMetricsReporterSampler | PR-1368 Automate leadership concurrency adjustment based on broker metrics | PR-1538 Enable partition metric collection to a configured topic during ongoing execution |
| PR-1234 Ensure consistent rackID during topic_configuration operations to avoid NPE | PR-1369 Enable Cruise Control to collect metrics from low traffic clusters by default | PR-1559 Add a ReplicaMovementStrategy that prioritizes (At/Under)MinISR partitions with offline replicas |
| PR-1241 Add support to retrieve capacity only via load endpoint | PR-1391 Calculate balance lower bound for resource distribution lower bound with low utilization threshold | PR-1589 Add parameters to relevant endpoints to control the speed of proposal generation |
| PR-1246 Enable Kafka port retrieval from listeners config | PR-1401 Honor webserver.api.urlprefix config | PR-1593 Set the default proposal generation speed to fast mode |
| PR-1255 Enable broker metric collection during ongoing executions | PR-1407 Provide idempotency support to handle duplicate maintenance events | PR-1599 Add a config to enable/disable provisioner |
| PR-1256 Prevent Cruise Control from mistakenly believe that there is an ongoing execution | PR-1409 [ccclient-1.1.2] Add topic parameter in cruise-control-client to query the kafka_cluster_state endpoint | PR-1604 Move broker failure detector away from using zNode-based failed broker persistence |
| PR-1282 Let executor substate show information on process before starting an execution | PR-1410 Add a sensor to emit topic count in cluster | PR-1608 Disable bootstrap endpoint in non-developer_mode |
| PR-1287 Prevent concurrent execution request from corrupting ongoing execution state | PR-1412 Enable detecting and fixing maintenance events during ongoing executions | PR-1622 Enable users to monitor the min.insync.replicas of all topics |
| PR-1289 Automate replica reassignment concurrency adjustment based on broker metrics | PR-1418 Add missing population of anomaly details for maintenance events | PR-1635 Add config to skip rack-awareness check while self-healing RF anomalies |
| PR-1294 Add sensors to emit concurrency caps for partition and leadership reassignments | PR-1419 Handle non-existent topic while setting/removing throttled replicas for a topic | PR-1636 Add capability to stop executions not started by Cruise Control |
| PR-1297 Make timeout for listing partition reassignments configurable along with a retry logic | PR-1433 Handle missing listeners config in CruiseControlMetricsReporter | PR-1637 Detect RF violations for topics having targetReplicationFactor with topicReplicationFactorMargin violation |

| | | |
|---|---|---|
| PR-1646 Setup ProvisionerState for rightsizing clusters | PR-1681 Handle metrics reporter exceptions while getting CPU metric | |

# What's New in Apache HBase

Learn about the new features of HBase in Cloudera Runtime 7.2.14.

## COD supports HBase 2.4.6

For smooth and better functionality, COD is updated to support HBase version 2.4.6. You need to upgrade the HBase client for seamless connectivity.

# What's New in Apache Impala

Learn about the new features of Impala in Cloudera Runtime 7.2.14.

## Priority Based Scratch Directory Selection

In this release, you have an option to configure the priority of the scratch directories based on your storage system configuration. The scratch directories will be selected for spilling based on how you configure the priorities of the directories and if you provide the same priority for multiple directories then the directories will be selected in a round robin fashion.

# What's New in Apache Kafka

Learn about the new features of Apache Kafka in Cloudera Runtime 7.2.14.

## Rebase on Kafka 2.8.0

Kafka shipped with this version of Cloudera Runtime is based on Apache Kafka 2.8.0. For more information, see the following upstream resources:

Apache Kafka Notable Changes:

- 2.6.0
- 2.7.0
- 2.8.0

Apache Kafka Release Notes:

- 2.6.0
- 2.7.0
- 2.8.0

## Kafka broker rolling restart checks

Cloudera Manager can now be configured to perform different types of checks on the Kafka brokers during a rolling restart. Using these checks can ensure that the brokers remain healthy during and after a rolling restart. As a result of this change, Kafka rolling restarts may take longer than in previous versions. This is true even if you disable the rolling restart checks. For more information, see Rolling restart checks.

## Http Metrics Report Exclude Filter introduced for Kafka

A new property, Http Metrics Report Exclude Filter (kafka.http.metrics.reporter.exclude.filter), is introduced for the Kafka service. This property can be used to specify a regular expression that is used to filter metrics. Any metric matching the specified regular expression is not reported by Cloudera Manager. As a result, these metrics are also not displayed in SMM. Use JMX metric names when configuring this property.

### Bootstrap servers are automatically configured for Kafka Connect

The Bootstrap Servers property of the Kafka Connect role is now automatically configured to include the bootstrap servers of its co-located Kafka brokers. This is only done if the property is left empty (default). You can provide custom value for this property if you want to override the default host:port pairs that Kafka Connect uses when it establishes a connection with the Kafka brokers.

### Kafka Connect Ranger Authorizer

A Ranger plugin is introduced for Kafka Connect that implements the Authorizer interface. A new service type is now also introduced in Ranger called kafka-connect. By default it includes the cm_kafka_connect resource-based service which includes policies that provide default access. The default resource-based service that is created for Kafka Connect can be configured using the 'Ranger service' name for the Kafka Connect service (ranger_plugin_kafka_connect_service_name) Kafka service property.

### Kafka Connect in DataHub [Technical Preview]

Kafka Connect can now be provisioned in CDP Public Cloud with Data Hub. The default Streams Messaging cluster definitions are updated to include Kafka Connect. For more information, see Streams Messaging cluster layout, Creating your first Streams Messaging cluster, and Scaling Kafka Connect.

### Stateless NiFi Source and Sink [Technical Preview]

The Stateless NiFi Source and Sink connectors enable you to run NiFi dataflows within Kafka Connect. Using these connectors can grant you access to a number of NiFi features without having the need to deploy or maintain NiFi on your cluster. For more information on the connectors, best practices on building dataflows to use with these connectors, as well as information on how to deploy the connectors, see Stateless NiFi Source and Sink.

### New Cloudera developed Kafka Connect connectors [Technical Preview]

In addition to the introduction of the Stateless NiFi Source and Sink, 12 new Cloudera developed connectors are available for use with Kafka Connect. These are powered by the Stateless NiFi engine and run Cloudera developed dataflows. They provide an out-of-the box solution for some of the most common use cases for moving data in or out of Kafka. For more information, see  Connectors in the Kafka Connect documentation.

### Kafka multiple Availability Zone support [Technical Preview]

Kafka can now be deployed in multiple Availability Zones in CDP Public Cloud. When using the multi Availability Zone feature, CDP ensures that Kafka replicates partitions across brokers in different availability zones. For more information, see Deploying CDP in multiple AWS availability zones.

## What's New in Apache Kudu

Learn about the new features of Kudu in Cloudera Runtime 7.2.14.

### Hive Metastore integration

Kudu can integrate its own catalog with the Hive Metastore (HMS). The HMS is the de-facto standard catalog and metadata provider in the Hadoop ecosystem. When the HMS integration is enabled, Kudu tables can be discovered and used by external HMS-aware tools, even if they are not otherwise aware of, or integrated with Kudu. Kudu supports table comments directly on Kudu tables which are automatically synchronized when the Hive Metastore integration is enabled. These comments can be added at table creation time and changed via table alteration. For more information, see Using Hive Metastore with Apache Kudu.

### Server startup progress page

A server startup progress page that shows the progress of the startup is implemented. The startup progress is broken down into the following steps:

1. Initializing
2. Reading Filesystem
   a. Reading instance metadata files
   b. Opening container files
3. Bootstrapping and opening the tables
4. Starting RPC server

The following metrics are exposed using milliseconds as the unit of time:

- log_block_manager_total_containers_startup : total containers present
- log_block_manager_processed_containers_startup : count of containers opened/processed until the requested instant of time
- log_block_manager_containers_processing_time_startup : time elapsed for opening the containers. If the containers are not yet opened, the time elapsed so far is provided.
- tablets_num_total_startup : total tablets present
- tablets_num_opened_startup : count of tablets opened/processed until the requested instant of time

tablets_opening_time_startup : time elapsed for opening the tablets. If the tablets are not yet opened, the time elapsed so far is provided.

### Improvements

- New metrics for amount of available space: Two new metrics wal_dir_space_available_bytes and data_dirs_space_available_bytes are introduced, indicating the amount of free space available in the WAL directory and collectively across data directories, respectively.
- New -list_statistics flag for 'kudu table list': The -list_statistics flag helps to list table's statistics such as tablet number, replica number, and record number.
- Support to dump and edit pbc in JSON pretty format: The kudu pbc edit and pbc dump tools are enhanced to dump PBC content in JSON pretty format by adding --json_pretty flag.
- Kudu version information can be checked when installing or upgrading using `curl http://x.x.x.x:8050/version`. It directly returns json information including the version.
- KUDU-3311: SysCatalogTable:Load allows cluster startup if the one master has two additional entries in --master_addresses.

## What's New in Apache Phoenix

Learn about the new features of HBase in Cloudera Runtime 7.2.14.

### COD supports HBase 2.4.6

For smooth and better functionality, Phoenix is now using HBase version 2.4.6.

## What's New in Schema Registry

Learn about the new features of Schema Registry in Cloudera Runtime 7.2.14.

**Support added for JSON schemas in Schema Registry**

The JSON type schema format is now supported.

Earlier, only Avro type schema format was supported out of the box.

For more information, see *Adding a new schema*.

**Cloudera Manager supports rolling restarts of HA enabled Schema Registry**

Schema Registry service can now be rolling restarted using Cloudera Manager.

**Import tool for Schema Registry schemas**

Schemas stored in Schema Registry can be exported to a JSON file. The exported JSON file can then be imported into another Schema Registry database. During an import, SchemaMetadata, SchemaBranch, and SchemaVersion objects are put into the database. These objects retain their ID as well as a number of other properties that are available in the JSON file used for import. This way, serializing and deserializing protocols can continue to function without any change and Schema Registry clients can seamlessly switch between different Schema Registry instances. Both import and export operations are done using the Schema Registry API.

# What's New in Cloudera Search

Learn about the new features of Cloudera Search in Cloudera Runtime 7.2.14.

## solr-client-socket-timeout cli option for HBase indexer

The --solr-client-socket-timeout optional argument overwrites the default 600000 millisecond (10 minute) socket timeout in HBase indexer 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).

# What's New in Apache Spark

Learn about the new features of Spark in Cloudera Runtime 7.2.14.

## 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 7.2.14.

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 7.2.14.

**Reactive Lineage fetching from Kafka producer cache**

You can now visualize the lineage between producers and consumers in SMM. Lineage information helps you to understand how the message is moving from a producer to a consumer group and which topics or partitions are part of that flow. Lineage between clients and topics or partitions are now shown using the new lineage endpoints. For more information, see Monitoring lineage information.

**New endpoint added to fetch lineage information for a topic**

The /api/v1/admin/lineage/partitions/{topic} endpoint used to fetch which producers have produced into the queried topic, and which consumerGroup's members have consumed from it. Now when you click on a topic to fetch the lineage on the UI, this endpoint is used.

**New endpoint added to fetch lineage information for a consumerGroup**

The /api/v1/admin/lineage/consumerGroups/{consumerGroupId} endpoint used to fetch which topics the members of that consumerGroup have consumed from, and also what producers have produced into those topics. Now when you click on a group on the UI to fetch the lineage, this endpoint is used.

**New endpoint added to fetch lineage information for a topicPartition**

The /api/v1/admin/lineage/partitions/{topic}/{partition} endpoint is used to fetch which producers have produced into that queried topicPartition and which consumerGroup members have consumed from that topicPartition. Now when you click on a topicPartition to fetch the lineage on the UI, this endpoint is used.

**New endpoint added to fetch lineage information for a producer**

The /api/v1/admin/lineage/lineage/producers/{producerId} endpoint is used to fetch which topics the queried producer has produced into, and which consumerGroups members have consumed from those topics. Now when you click on a producer to get the lineage on the UI, this endpoint is used.

**On selecting the partition on Overview page, the new lineage endpoint should be called**

Lineage between clients and topics or partitions are now shown using the new lineage endpoints. Remember that when checking the lineage (connected clients) for a TopicPartition, only the recently connected clients will be shown.

**Support added for multiple replication targets in SMM**

SMM now supports SRM replication flows targeting remote clusters making use of the new v2 SRM APIs.

Remote Replication flows available under the “/api/v2/admin/replication-stats” APIs. The UI is now configured to make use of these new APIs.

**Introduced multi-target replication monitoring support in alerts**

- SMM now adopted the new V2 SRM Service endpoints upon which alerting is based on.
- In SMM when configuring alerts for replications in the UI now source and target clusters can be defined, as opposed to the previous configuring panel, where only the source cluster could be defined (since the target cluster was fixed to be the colocated Kafka cluster).
- Old alerts will still function, however editing them can only be done using the new format, where source and target clusters have to be defined.
- IMPORTANT: For alerts involving remote SRM cluster queries set the execution interval to at the very least a minute (preferably more).

**SMM authenticates to SRM Service**

SMM now automatically configures Basic Authentication when connecting to SRM and the service dependency based auto-configuration is in use.

For manual SRM connectivity configurations, Basic Auth configurations were added (Streams Replication Manager Basic Authentication, Streams Replication Manager Basic Authentication Username, Streams Replication Manager Basic Authentication Password).

**SMM Cache-Control is part of default SMM REST Server API's responses' headers**

The new SMM configuration named cache.control.http.response.header.value allows to configure the Cache-Control header's value for certain endpoints. Configure it in the following key-value like fashion:

- The key is the path prefix to the endpoints where the Cache-Control header should be added.
- The value is the value of Cache-Control header.

In order to turn off functionalities provided by the Cache-Control header just delete the entries, or set the value to no-store.

**Added helper tooltips to SMM UI**

SMM now provides more informative tooltips (hover over the table headers and labels) for most of its elements in the web UI.

**Removed Consumer Rate graphs**

The lag rate graph is removed from the UI.

The lag rate values are removed from the /api/v1/admin/metrics/aggregated/groups/{groupName} and /api/v1/admin/metrics/aggregated/groups endpoints.

**SMM is automatically integrated with co-located Kafka Connect**

To monitor and manage Kafka Connect in SMM, a number of SMM service properties must be configured. These are the following:

- Kafka Connect Host
- Kafka Connect Port
- Kafka Connect Protocol

From now on, these properties are automatically configured if Kafka Connect Host is left empty (default). This means that the SMM service automatically configures itself to connect to its co-located Kafka Connect instance. You can provide custom values for all properties if you want to override the defaults.

## What's New in Streams Replication Manager

Learn about the new features of Streams Replication Manager in Cloudera Runtime 7.2.14.

### SRM Driver monitoring using Cloudera Manager

Cloudera Manager’s ability to monitor the SRM Driver, its replications, and the overall health of SRM is improved. Most notably, the health status of SRM is based on the health of the network and the availability of replication sources and targets. As a result of this improvement, two new metrics, a new health test, and several new configuration properties are introduced for the SRM Driver in Cloudera Manager.

**New metrics and health test**

The new metrics are as follows:

- SRM Driver Distributed Herder Status (streams_replication_manager_distributed_herder_status)

- Aggregated Status Code of SRM Driver Replication Flows (streams_replication_manager_aggregated_herder_status)

The distributed metric describes the status of individual replications. The aggregate metric provides the aggregate status of all replications.

The new health test is called DISTRIBUTED_HERDER_STATUS. This health test is based on the aggregate metric and provides information on the overall status of SRM and its replications.

**New properties**

The new monitoring related properties are as follows:

- Path for driver plugins (plugin.path)
- Enable HTTP(S) Metrics Reporter (mm.metrics.servlet.enable)
- SSL Encryption for the Metrics Reporter (metrics.jetty.server.ssl.enabled)
- HTTP Metrics Reporter Port (metrics.jetty.server.port)
- HTTPS Metrics Reporter Port (metrics.jetty.server.secureport)
- Enable Basic Authentication for Metrics Reporter (metrics.jetty.server.authentication.enabled)
- Metrics Reporter User Name (metrics.jetty.server.auth.username)
- Metrics Reporter Password (metrics.jetty.server.auth.password)

For more information, see Cloudera Manager Configuration Properties Reference.

## SRM Driver now automatically retries setting up replications for unavailable target Kafka clusters

Previously, if any of the Kafka clusters that were targeted by the SRM Driver were unavailable at startup, the SRM Driver stopped. As a result of an improvement, the SRM Driver now instead sets up replications for all target Kafka clusters that are available and continuously retries to set up replication for unavailable clusters. Retry behaviour is configurable in Cloudera Manager. The new properties related to retry behaviour are as follows:

- Retry Count for SRM Driver (mm.replication.restart.count)
- Retry Delay for SRM Driver (mm.replication.restart.delay.ms)

For more information see, Cloudera Manager Configuration Properties Reference or Configuring SRM Driver retry behaviour.

## Disabled replications can now be fully deactivated by configuring heartbeat emission

As a result of the rebase to Kafka 2.8 (KAFKA-10710), an improvement is introduced in connection with heartbeat emission. From now on, you can fine tune your deployment and fully deactivate any unnecessary replications that are set up by default by configuring heartbeat emission. This can help with minimizing any performance overhead caused by unnecessary replications.

To support this change, an improvement was made for the SRM service in Cloudera Manager. A dedicated configuration property, Enable Heartbeats, is introduced. You can use this property to configure emit.heartbeats.enabled on a global level directly in Cloudera Manager. Replication level overrides are still supported. This can be done by adding emit.heartbeats.enabled with a valid replication prefix to Streams Replication Manager's Replication Configs. For more information on configuring heartbeat emission, see Configuring SRM Driver heartbeat emission.

## IdentityReplicationPolicy now available

**Warning:** The IdentityReplicationPolicy does not detect cycles. As a result, using this replication policy is only viable in deployments where the replication setup is acyclic. If your replication setup is not acyclic, using this replication policy might result in records being replicated in an infinite loop between clusters. Additionally, monitoring replications with the SRM Service is not possible when this policy is in use.

The version of Apache Kafka shipped with this release of Cloudera Runtime includes KAFKA-9726. As a result, the IdentityReplicationPolicy is available for use with Streams Replication Manager. This replication policy does not rename remote (replicated) topics. Streams Replication Manager can be configured to use this replication policy by adding the following entry to Streams Replication Manager's Replication Configs:

```
replication.policy.class=org.apache.kafka.connect.mirror.IdentityReplication
Policy
```

For more information, see KAFKA-9726.

### SRM configuration properties can be configured globally for Connect workers and Connect connectors

The SRM Driver now accepts configuration properties prefixed with the workers. and connectors. prefixes. Configuration properties added to Streams Replication Manager's Replication Configs that use these prefixes are applied globally to all Connect workers or Connect connectors that the SRM Driver creates. For more information regarding the prefixes, see Understanding SRM properties, their configuration and hierarchy. For more information on Connect workers and connectors, see Streams Replication Manager Architecture.

### SRM Service Basic Authentication support

The SRM Service can now be secured using Basic Authentication. Once Basic Authentication is set up and enabled, the REST API of the SRM Service becomes secured. Any clients or services that connect to the REST API will be required to present valid credentials for access. Configuration is done in Cloudera Manager using SRM configuration properties and external accounts.

For more information, see Configuring Basic Authentication for the SRM Service.

### SRM automatically creates a Basic Authentication credential for co-located services

SRM automatically creates a Basic Authentication credential for co-located services (users can change the credentials using SRM Service Co-Located Service Username and SRM Service Co-Located Service User Password). When Basic Authentication is enabled, this user is automatically accepted by the SRM Service. For more information, see Configuring Basic Authentication for the SRM Service.

### SRM Service Remote Querying no longer in technical preview

SRM Service Remote Querying was introduced in a previous release of Cloudera Runtime as a technical preview feature. Starting with this release, Remote Querying is ready for use in production environments. This is the result of Basic Authentication being introduced for the SRM Service and SMM supporting multi-target alerting.

For more information on Remote Querying, see Remote Querying and Configuring Remote Querying. For more information on how to set up Basic Authentication for Remote Querying, see Configuring Basic Authentication for Remote Querying.

### The SRM Driver can now write the origin offset into the record header

SRM now supports a diagnostic feature in which the source offset of the replicated records are written into the headers. The feature can be turned on by setting copy.source.offset.in.header.enabled to true. When enabled, the source offset is written into a header named mm2-source-offset in binary format. The schema of the header payload is available in the connect:mirror-client package, the class name is org.apache.kafka.connect.mirror.SourceOffsets. This feature is only recommended for diagnostic purposes, as the header change increases the size of the replica topic.

## What's New in Apache Hadoop YARN

Learn about the new features of Hadoop YARN in Cloudera Runtime 7.2.14.

### UI update for Dynamic Queue Scheduling

The YARN Queue Manager UI for the Dynamic Queue Scheduling feature has been revised. The Set Schedule step is now called Specify Schedule and it display all supported scheduling options on a single page. Previously, the available options were organized into multiple tabs.

For more information about the Dynamic Queue Scheduling feature, see Dynamic Queue Scheduling [Technical Preview].

### Editing placement rules

Support to edit previously created placement rules was added. For more information, see Edit placement rules.

### Setting Maximum Parallel Application Limits

You can set the maximum number of applications that can run at the same time in a cluster. You can set parallel application limits for all queues, a specific queue, all users, and at the user level.

For more information, see Setting Maximum Parallel Application limits.

### 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 7.2.14.

- Apache Atlas
- Data Analytics Studio
- Apache Hadoop HDFS
- Apache Hive
- Hue
- Apache Knox
- Apache Oozie
- Apache Ranger
- Apache ZooKeeper

# Fixed Issues In Cloudera Runtime 7.2.14

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.

## Fixed Issues in Atlas

Review the list of Apache Atlas issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-29335: Random NPE when retrieving tasks.**

Handle null pointer exception when retrieving tasks using the admin/tasks endpoint.

**CDPD-29663: Error while connecting topic with schema in Atlas**

The error occurred when the Schema Registry tried to make a relationship in Apache Atlas between a schema and a non-existent corresponding topic. When the error occurred, an Atlas integration related exception showed up in the Schema Registry logs and a DB entry related to the Atlas task was left in a failed state. The error is safe to ignore.

### Apache patch information

- ATLAS-4431
- ATLAS-4548
- ATLAS-4454

## Fixed Issues in Avro

Review the list of Avro issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-24010: Upgrade to velocity 2.3 due to CVE-2020-13936.**

The Velocity dependency has been updated to 2.3. so CVE-2020-13936 is fixed in Avro. This issue is now resolved.

### Apache patch information

Apache patches in this release. These patches do not have an associated Cloudera bug ID.

- AVRO-2228

## Fixed Issues in Cloud Connectors

Review the list of Cloud Connectors issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-12425: support s3 client side encryption.**

The s3a connector now supports S3-CSE client side encryption. This issue is resolved.

**CDPD-33981: S3A auditing has been disabled to avoid memory leaks in hive. To re-enable s3a auditing, set fs.s3a.audit.enabled to true.**

This issue is resolved.

### Apache patch information

- HADOOP-18094
- HADOOP-13887

## Fixed issues in Cruise Control

There are no fixed issues for Cruise Control in Cloudera Runtime 7.2.14.

**CDPD-33535: Upgrading Logredactor version**

The Logredactor version is upgraded to 2.0.13 version to fix CVE-2021-44228 issues.

## Fixed issues in Data Analytics Studio

There are no fixed issues for Data Analytics Studio in Cloudera Runtime 7.2.14.

## Fixed Issues in Apache Hadoop

Review the list of Hadoop issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-20357: WARN concurrent.ExecutorHelper: Thread (Thread[GetFileInfo #0,5,main]) interrupted.**

This issue is resolved.

### Apache Patch Information

- HADOOP-18065

## Fixed Issues in HBase

Review the list of HBase issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-22120: Backport HBASE-25006 and HBASE-22618 - Make the cost functions optional for StochastoicBalancer.**

Added the possibility to load custom cost functions. Optional for StochasticLoadBalancer.

**CDPD-32297: HBase master fails to active on AWS.**

Reverting stream closure on hbck lock file to fix issue of HMaster becoming active on AWS. This issue is resolved.

### Apache Patch Information

- HBASE-25006
- HBASE-26512

## Fixed Issues in HDFS

There are no fixed issues for HDFS in Cloudera Runtime 7.2.14.

## Fixed Issues in Apache Hive

Review the list of Hive issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-29779: Various hive query failures: HiveAccessControlException: HivePrivilegeObjects are not available to authorize this command.**

Reverted CDPD-30250 which is the cause of this issue. This issue is resolved.

### Apache Patch Information

- TEZ-4238

## Fixed Issues in Hive Warehouse Connector

There are no fixed issues for HWC in Cloudera Runtime 7.2.14.

## Fixed Issues in Hue

There are no fixed issues for Hue in Cloudera Runtime 7.2.14.

# Fixed Issues in Apache Impala

There are no fixed issues for Impala in Cloudera Runtime 7.2.14.

# Fixed Issues in Apache Kafka

Review the list of Apache Kafka issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-27780: IdentityReplicationPolicy for MM2 to mimic MM1 (KAFKA-9726 backport)**

**Warning:** The IdentityReplicationPolicy does not detect cycles. As a result, using this replication policy is only viable in deployments where the replication setup is acyclic. If your replication setup is not acyclic, using this replication policy might result in records being replicated in an infinite loop between clusters. Additionally, monitoring replications with the SRM Service is not possible when this policy is in use.

This is a backported improvement that introduces a new replication policy called IdentityReplicationPolicy. This replication policy does not rename remote (replicated) topics. As a result of this backport, the IdentityReplicationPolicy is now available for use with Streams Replication Manager. For more information, see KAFKA-9726.

**OPSAPS-61697: Kafka broker IDs are overridden when importing a cluster template**

Importing a cluster template to a new cluster in Cloudera Manager no longer overrides Kafka broker IDs (broker.id) or other role specific unique identifiers if the unique identifiers are already set in the template.

# Fixed Issues in Apache Knox

Review the list of Knox issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-30711: The Token Integration link on CP UI is available for all customers going forward..**

This issue is resolved.

**CDPD-19654: The Console URL on the web UI is not a Knox URL.**

The Console URL for Yarn on the Oozie UI will be a Knox URL when the Oozie UI is accessed through Knox. This issue is resolved.

**CDPD-28900: Authentication with IDBroker failed in dex post test runs for DL HA Terminate flow.**

Make sure correct SNI header is added after failover. This issue is resolved.

**CDPD-30731: Knox rewrite not happening for new v2 SMM endpoints**

Fixed SMM not showing Replications tab on Public Cloud when Knox is enabled. This issue is resolved.

## Apache patch information

Apache patches in this release.

- KNOX-2675
- KNOX-2696
- KNOX-2698

# Fixed Issues in Apache Kudu

Review the list of Apache Kudu issues that are resolved in Cloudera Runtime 7.2.14.

**Validation for extra configuration properties for a table**

A new validation logic is added. It is performed on the strings representing the keys of extra configuration properties for a Table. As a result an error is returned on attempt to set unsupported extra configuration properties for a table.

**Custom hash schema criterion in KuduTableCreator::Create()**

This patch fixes a bug in KuduTableCreator::Create(). The criterion for the presence of custom hash schemas per range flag accounts when the table-wide hash schema is non-empty, but there is a range partition with empty hash schema.

**KUDU-75: Allow RPC proxies to take HostPort and do DNS resolution inline with calls**

Kudu will now automatically re-resolve addresses in cases where a tablet server address is no longer valid. This previously manifested itself as a crash when DNS resolution between servers failed.

**KUDU-1938: Support for VARCHAR type**

It's now possible to create Kudu tables with VARCHAR columns using impala-shell since KUDU-1938 and IMPALA-5092 have been addressed.

**KUDU-3295: kudu client cannot read if one of tablet replica is unhealthy**

When the tserver is shutdown and its DNS cannot be resolved, it is still possible for the client to receive the replicas including this tserver. As a result, discoverTablets sees the servers size not equal to replica size because the replica does not verify server connectivity. This patch fixes this issue by ignoring replicas whose server cannot be resolved.

**KUDU-3334: Clarify whether it's possible to hide libprotobuf symbols exported from libkudu_client**

Conflicts happened during static initialization of symbols related to protobuf extensions in an application using libkudu_client and also linking in libprotobuf on its own.This patch adds the --exclude-libs flag for the GNU linker to completely hide all symbols inherited by libkudu_client from libprotobuf.

**KUDU-3341: Catalog Manager should stop retrying DeleteTablet when receive WRONG_SERVER_UUID error**

When the calatog_manager tries to delete a tablet but gets WRONG_SERVER_UUID error it will not retry this task because the served uuid can only be corrected by restarting the tablet server.

## Fixed Issues in Apache Oozie

Review the list of Oozie issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-21874: Oozie should be able to handle file-system credentials coming from various places in the workflow.xml.**

This issue is resolved.

**CDPD-30584: The Oozie client will display the full stacktrace in case it fails to connect to the Oozie server.**

This issue is resolved.

**CDPD-30045: Oozie's ZooKeeper related codebase now has extra trace level logging entries so in case of any ZooKeeper related issues these can be enabled. Affected classes: org.apache.oozie.command.XCommand org.apache.oozie.service.ZKLocksService**

This issue is resolved.

**OPSAPS-57546: All form factors. When the Knox gateway is available on the cluster and it’s discovery is enabled for Oozie then the Web UI link of Oozie through Knox will appear among the direct links.**

This issue is resolved.

**OPSAPS-62019: Enabled thread level logging for Oozie server by adding "%t" in log4j properties in Oozie configs.**

This issue is resolved.

**CDPD-28768: Actions are stuck while running in a Fork-Join workflow.**

This issue is resolved.

**CDPD-30426: Correct maximum wait time between database retry attempts property.**

The oozie-default.xml had a default value for controlling the JPA retry wait time. The name of the property in the oozie-default.xml was: oozie.service.JPAService.retry.maximum-wait-time.ms. However, the Oozie server was looking for a property named oozie.service.JPAService.maximum-wait-time.ms. This is now fixed and the oozie-default.xml and the Oozie server are now in sync. In case you put an override as a safety-valve with a name oozie.service.JPAService.maximum-wait-time.ms, then Oozie will prefer that.

**OPSAPS-61115: Implement a new checkbox for Oozie to disable the Oozie UI**

A new checkbox was implemented on the Oozie configuration page which can be used to turn off the Oozie UI completely. Meaning none of the Oozie UI resources will be served and thus if you are worried about JQuery vulnerabilities, etc. which cannot be fixed short term, you can use this feature to get rid of these by not exposing these at all.

**CDPD-27661: Implement a way to disable the Oozie UI**

There is a new property named "oozie.ui.enabled" users can set for Oozie with a value true of false. By default it's set to true. When set to false, the Oozie UI will be complete disabled, the Oozie server will not even expose the UI resources, hence this can be a workaround for the JQuery vulnerabilities.

**CDPD-19654: The Console URL on the web UI is not a Knox URL.**

The Console URL for Yarn on the Oozie UI will be a Knox URL when the Oozie UI is accessed through Knox.

**CDPD-30487: Improve getActionData in Oozie.**

The Yarn application launched by Oozie creates a file for storing action related information which will be read by the Oozie server. In case the Yarn application is able to create this file, but won't be able to put the necessary data into it and so the file remains empty, the Oozie server will fail to parse this file. This error is now fixed.

### Apache patch information

OOZIE-3422

## Fixed Issues in Phoenix

There are no fixed issues for Phoenix in Cloudera Runtime 7.2.14.

## Fixed Issues in Parquet

There are no fixed issues for Parquet in Cloudera Runtime 7.2.14.

### Apache Patch Information

- PARQUET-2094

## Fixed Issues in Apache Ranger

Review the list of Ranger issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-28752: Provided sorting on specific columns on policy and on audit listing page.**

This issue is resolved.

**CDPD-31371: Added RAZ policy for RAZ Azure datalake auto backup/ restore.**

This issue is resolved.

**CDPD-17304: Make Ranger Solr audit collection storage configurable via CM.**

Made changes to use the custom config set zip file to create the config-set which is used by Ranger Admin start script to create the collection, previously config-set location was fixed. This issue is resolved.

**CDPD-21398: Ranger end to use the first PrivateKeyEntry in a keystore**

After the fix Ranger Admin now allows user to customise the private keystore instead of using the default value. This issue is resolved.

**CDPD-27138: Enhance Ranger admin REST Client to use cookie for policy, tag and role download.**

Apache Ranger REST Client to download policies, tags and roles from Ranger admin will use cookie session. Earlier each of the plugin has to do kerberos login to get a TGT to download policy, tags and roles. With this feature Session Cookie is enabled by default in RangerAdminClient and it will be used instead of hitting KDC for TGT for validating the user. This improve performance as well and reduce the load on KDC.

**CDPD-29866: Ranger admin REST Client is not using cookie by default.**

This issue was caused because a feature build in https://jira.cloudera.com/browse/CDPD-27138 was ported to 7.2.12.0 and this JIRA resolved it.

### Apache Patch Information

- RANGER-3334
- RANGER-3276
- RANGER-3519
- RANGER-3535
- RANGER-2634
- RANGER-3490
- RANGER-3512
- RANGER-3520
- RANGER-3504
- RANGER-3503
- RANGER-3481
- RANGER-3518
- RANGER-3515
- RANGER-3514
- RANGER-3505
- RANGER-3533
- RANGER-3023

## Fixed Issues in Schema Registry

Review the list of Schema Registry issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-32192: First start failed for SR, with oracle DB, migration failed at CREATE TABLE "atlas_events"**

Fixed v009__create_registry_audit.sql to have create index refer to the lower case "atlas_events" object (the table).

Made the script re-runnable since the table was already created where the script had already run.

**CDPD-31907: Schema Registry REST API endpoint does not show SchemaBranches**

Schema Registry’s /api/v1/schemaregistry/schemas/aggregated REST API endpoint shows SchemaBranches without SchemaVersions.

**CDPD-30996: SR does not create new SchemaMetadata with given ID**

In DefaultSchemaRegistry class, addSchemaMetadata (Supplier<Long> id, SchemaMetadata schemaMetadata, boolean throwErrorIfExists) does not look for the given ID, but the next available ID.

**CDPD-29700: Hide Compatibility list in the website**

When the schema type is JSON, then the Compatibility field will be hidden in the website.

**CDPD-29663: Error while connecting topic with schema in Atlas**

When Schema Registry tries to make a relationship in Atlas between a schema and a non-existent corresponding topic, an error occurs.

**CDPQE-11299: Importing schemas in Confluent format might fail**

Fixed the issue where importing from the Confluent Schema Registry fails intermittently.

## Fixed Issues in Cloudera Search

Review the list of Cloudera Search issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-29853: Solr 8 now works in secure environments without sending OPTIONS requests during internal communication**

This fix eliminates unnecessary OPTIONS requests in internal Solr communication, improving performance.

**CDPD-23110: Hiveserver2/HMS hung because of the LeaseRenewer thread is waiting to get the kerberos ticket via System.in**

Solr client does not overwrite ‘null’ value of  javax.security.auth.useSubjectCredsOnly parameter, it only throws a warning during connection setup:

```
System Property: javax.security.auth.useSubjectCredsOnly set to:
 [true|null] not false.
SPNego authentication may not be successful.
```

This may cause issues when connecting Solr to custom applications. To prevent this, you need to set

-D   javax.security.auth.useSubjectCredsOnly=false

in the JVM configuration of those applications. Cloudera has implemented this change in MapReduceIndexer, CrunchIndexer, Spark-Solr, Hive-Solr connector, and Atlas.

## Fixed Issues in Apache Solr

There are no fixed issues for Solr in Cloudera Runtime 7.2.14.

**Solr 8 now works in secure environments without sending OPTIONS requests during internal communication**

This fix eliminates unnecessary OPTIONS requests in internal Solr communication, improving performance.

### Apache patch information

None

## Fixed Issues in Spark

Review the list of Spark issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-28196: RM UI redirect link to the Spark3 History Server not working.**

Spark 3 History Server link Resource Manager UI works again. This issue is resolved.

**CDPD-30201: Fix Event Timeline in SparkUI**

Fixed "Event Timeline" expansion in SparkUI. From 7.2.14 onwards, clicking "Event Timeline" in SparkUI works as expected. This issue is resolved.

### Apache patch information

None

## Fixed Issues in Apache Sqoop

Review the list of Sqoop issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-30696: Sqoop Hive imports should preserve KRB5CCNAME by default.**

When Sqoop runs the Hive import in a new process it will preserve some of the environment variables of the parent process. From now on Sqoop will also automatically preserve the KRB5CCNAME environment variable, so in case of a customer Kerberos ticket file is used, the underlying beeline process will also be aware of it. This issue is resolved.

### Apache patch information

No additional Apache patches.

## Fixed Issues in Streams Messaging Manager

Review the list of Streams Messaging Manager issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-30745: Broken link to the TopicDetailPage**

When a topic is selected in the Replications tab, where the topic in question is not present on the local Kafka Cluster monitored by the current SMM instance, link breaks to the TopicDetailPage.

**CDPD-30731: Knox rewrite not happening for new v2 SMM endpoints**

SMM is not showing the Replications tab on Public Cloud when Knox is enabled.

**CDPD-30370: When TLS is enabled, SMM should connect to Schema Registry**

When TLS is enabled, SMM by default cannot properly connect to Schema Registry. As a result, the SMM Data Explorer shows errors when viewing Avro formatted data.

**CDPD-28002: The Cluster Replications tab is missing from the SMM UI**

SMM UI now correctly renders Replication workflows and metrics in Public Cloud.

**CDPD-24943: Long resource names (such as topic names, hostnames etc) are truncated on SMM UI**

Resource names in listings (such as topic name, host name, consumer name and so on ) are now overflowing to the next line with breaking text on any character, instead of remaining in 1 line with hidden overflow.

## Fixed Issues in Streams Replication Manager

Review the list of Streams Replication Manager issues that are resolved in Cloudera Runtime 7.2.14.

**OPSAPS-61814: Using the service dependency method to define Kerberos enabled co-located clusters is not supported**

When the Streams Replication Manager Co-located Kafka Cluster Alias configuration is used to auto-configure the connection to the co-located Kafka cluster, and Kerberos is enabled, the JAAS configuration is dynamically generated on each host. As a result, you can now use the service dependency method to define a Kerberos enabled co-located cluster.

**CDPD-31235: Negative consumer group lag when replicating groups through SRM**

SRM no longer tries to create a checkpoint or synchronize the group offset if there is no mapping available for the topic-partition in the offset-syncs topic.

## Fixed Issues in Apache YARN

Review the list of YARN issues that are resolved in Cloudera Runtime 7.2.14.

**COMPX-5252: QM does not support direct migration between weight and absolute resource allocation mode**

YARN Queue Manager UI did not support direct migration neither from weight resource allocation mode to absolute resource allocation mode, nor from absolute resource allocation mode to weight resource allocation mode. This issue is resolved.

**COMPX-6699: Revisit permission of /tmp/logs directory**

The permission of the /tmp/logs HDFS directory has been fixed. The user=yarn is now the owner of inode=/tmp/logs.

**COMPX-7493: YARN Tracking URL that is shown in the command line does not work when knox is enabled**

This issue is resolved.

**COMPX-7887: Fix CVE-2020-8908**

A temp directory creation vulnerability exists in all versions of Guava, allowing an attacker with access to the machine to potentially access data in a temporary directory created by the Guava API com.google.common.io.Files.createTempDir(). This issue is resolved.

**COMPX-8162: Maximum AM resource percentage value is not updated for dynamically created queues.**

This issue is resolved.

### Apache patch information

- MAPREDUCE-7307
- YARN-6091
- YARN-8148
- YARN-8659
- YARN-8864
- YARN-8984
- YARN-9011
- YARN-9290
- YARN-9584
- YARN-9601
- YARN-9640
- YARN-9642
- YARN-9714
- YARN-9728
- YARN-9956
- YARN-9993
- YARN-10364
- YARN-10393
- YARN-10438
- YARN-10467
- YARN-10501
- YARN-10555
- YARN-10649
- YARN-10651

- YARN-10701
- YARN-10703
- YARN-10720
- YARN-10934
- YARN-10980

## Fixed Issues in Zeppelin

Review the list of Zeppelin issues that are resolved in Cloudera Runtime 7.2.14.

### Apache patch information

Apache patches in this release.

CDPD-31506: Fix dependency convergence for jetty-util-ajax

ENGESC-10231: ZEPPELIN-4952: Markdown interpreter can be used to store XSS in notebooks

ENGESC-10231: [ZEPPELIN-4311] Supporting new parser for markdown as pegdown parser

CDPD-30912: CDH: 7.1.8.0 build failed for zeppelin for all OS builds

CDPD-29838: Zeppelin notebook creation failure due to Hadoop ClassCastException

CDPD-28967: Zeppelin - Upgrade commons-io to 2.8 due to CVE-2021-29425

CDPD-29175: Zeppelin - Upgrade jsoup to 1.14.2 due to CVE-2021-37714

CDPD-28924: Zeppelin - Upgrade to junit 4.13.2 due to CVE-2020-15250

CDPD-28834: Credentials file should get saved along with notebook-authorization.json and interpreter.json

CDPD-24981: Zeppelin notebook can not create table with jdbc phoenix interpreter

CDPD-23410: Zeppelin notebook_authorizations::test_only_owners_can_change_permissions test is failing

CDPD-22469: ZEPPELIN-5231: Livy Interpreter doesn't support Japanese Character - Encoding Issue

CDPD-17187: Zeppelin - Upgrade to Angular 1.8.0 due to CVEs

CDPD-20908: Remove log4j-slf4j-impl from JDBC/Hive interpreter

CDPD-19308: Zeppelin - Upgrade to slf4j 1.7.30

CDPD-19316: Zeppelin - Upgrade httpclient to 4.5.13+ / 5.0.3+ due to CVE-2020-13956

CDPD-20461: Zeppelin - Upgrade jackson to 2.10.5.1 or 2.11.0+ due to CVE-2020-25649

CDPD-20267: Zeppelin build failure on cdpd-master

CDPD-17471: [ZEPPELIN-5116]Accessing zeppelin via knox after knox logout should be redirected to knox login page

CDPD-17933: Zeppelin - Upgrade spring framework to 4.3.29.RELEASE+ due to CVE-2020-5421

CDPD-19243: Upgrade to Shiro 1.7.0 due to CVE-2020-17510

CDPD-18170: Zeppelin - Upgrade or remove auto-value due to shaded guava CVEs

CDPD-15497: Harmonize joda-time to version 2.10.6(cdpd harmonized)

CDPD-17543: Zeppelin UI is not comping due to Corrupted notebooks

CDPD-16197: Upgrade api-*-1.0.0.jar due to CVEs

CDPD-16096: Zeppelin - upgrade google-oauth-client to 1.31.0

CDPD-17017: Upgrade xercesImpl to to 2.12.0 due to CVE-2018-2799

CDPD-16845: Upgrade to Shiro 1.6.0 (CVE-2020-13933)

CDPD-16111: Upgrade jsoup-1.7.2 (CVE-2015-6748)

CDPD-16104: Upgrade postgresql JDBC driver to 42.2.16

CDPD-14569: [ZEPPELIN-4414]. Upgrade thrift to 0.13

CDPD-13378: Bumup version of Java Native Access (JNA)

CDPD-16114 Upgrade jackrabbit-webdav 1.5.2 due to CVE-2015-1833

ZEP-97: [ZEPPELIN-3690] display all column with the same name in ui-grid

CDPD-16115: Upgrade jgit due to CVE-2016-5725

CDPD-14614: Update Spring Framework for Zeppelin (CVE-2018-1270)

CDPD-11599: Update Quartz Enterprise Job Scheduler for Zeppelin (CVE-2019-13990)

CDPD-14580: Upgrade Scala for CVEs

BUG-124121: Password hashing not working in Zeppelin

CDPD-15628: Compilation faliure on dex-box

CDPD-12920: Upgrade nimbus-jose-jwt to 7.9

CDPD-14990: Upgrade libpam4j to 1.11 (CVE-2017-12197)

CDPD-11426: Ensure consistent usage of jackson to 2.10.3

CDPD-14579: remove org.reflections (CVE-2020-10683)

CDPD-14581: Update Spring Framework for Zeppelin in 7.2.1.0 (CVE-2018-1275)

CDPD-14369: [ZEPPELIN-4736] The use of SslContextFactory is deprecated

CDPD-11406: Include NOTICE and LICENSE files in component directories

CDPD-11301: Remove jackson and jersey-bundle

CDPD-11780: Zeppelin: Remove spark (and other interpreters that are not shipped) source dependencies

CDPD-11348: Update log4j to address CVE-2019-17571

CDPD-11501: Update Apache Shiro for Zeppelin to 1.5.3

CDPD-11571: Zeppelin build failure on cdpd-master due to perfect-scrollbar

CDPD-10187: Zeppelin - Incorrect version of jackson-mapper-asl in CDP

CDPD-9119: Zeppelin - Upgrade to Guava 28.1 to avoid CVE-2018-1023

CDPD-9030: Upgrade jackson-databind to version 2.9.10.3 [CVE-2020-8840]

CDPD-9454: [ZEPPELIN-4697] Zeppelin scheduler pings terracotta.org

CDPD-8163: Remove `org.spark-project.hive` dependency

CDPD-7789: Zeppelin - Upgrade to Jetty 9.4.26 to avoid CVEs

CDPD-7479: add hadoop-cloud-storage jar in Zeppelin

CDPD-3600: Sync Zeppelin with community latest version (0.8.2)

CDPD-2933: [ZEPPELIN-4272] Zeppelin fails to use s3a configured for zeppelin.notebook.dir

CDPD-1683: KerberosRealm roles should match with local file system, if nothing is specified

CDPD-2300: Initialize proxyuser with proper configuration

CDPD-1491: Zeppelin should support doAs

BUG-120595: [ZEPPELIN-4197] Upgrade Jackson to 2.9.9

BUG-120606: [ZEPPELIN-4187] Bump up version of Scala from 2.11.8 to 2.11.12 (#3378)

BUG-120605: [ZEPPELIN-4186] Bump up version of org.jsoup:jsoup (#3377)

BUG-120596: [ZEPPELIN-4188] Upgrade Jetty to 9.4.18.v20190429

BUG-120594: ZEPPELIN-4193 Upgrade Bouncy Castle bcpkix-jdk15on to 1.60

BUG-120593: [ZEPPELIN-4185] Upgrade Thrift to 0.12.0 (#3376)

CDPD-1009: ZEPPELIN-4168: Use secure URLs for Maven repositories (#3370)

CDPD-717: [Zeppelin 3792] Zeppelin SPNEGO support

ZEP-79: Disable fs.file.impl cache to ensure RawLocalFS is used

BUG-109581: [ZEPPELIN-3741] Do not clear "Authorization" header if Z-server is running behind proxy

BUG-106906: Add shiro-tools-hasher in Zeppelin

BUG-106297: JDBC interpreter log file is missing in zeppelin log directory

BUG-102172: Include Google Connector in Zeppelin

BUG-98604: Correct tutorial link should be added in interpreter page

BUG-100845: Remove livy2.pyspark3 interpreter on zeppelin side

BUG-114354: Fixes to make s3 storage work

BUG-114354: Change Zeppelin to use unshaded jars

BUG-103954: Exclude other dependencies in STS shaded JDBC driver to prevent conflict

BUG-103715: fix handshake_failure download

CDPD-10288: Zeppelin Notebook Initialisation fails with CNF error in RAZ Enabled Cluster

## Fixed Issues in Apache ZooKeeper

Review the list of ZooKeeper issues that are resolved in Cloudera Runtime 7.2.14.

**CDPD-25039: Prevent unnecessary client connection retry caused by slow SASL login**

This makes the ZooKeeper client connection / session initiation on kerberized clusters more stable.

### Apache Patch Information

- ZOOKEEPER-3590
- ZOOKEEPER-4275
- ZOOKEEPER-4367
- CURATOR-525

# Fixed Issues In Cloudera Runtime 7.2.14.3

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.3.

The following issues are resolved:

- HOTREQ-950 IDBroker client excessively adds SSL client config causing OOM issues
- HOTREQ-964 Release: HIVE-25574: Replace clob with varchar when storing creation metadata
- HOTREQ-1003 Optimisations on COD for ABFS support

# Fixed Issues In Cloudera Runtime 7.2.14.4

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.4.

The following issues are resolved:

- HOTREQ-1051 Query based compaction fails in Public Cloud
- HOTREQ-1036 Bug Fix for SPARK-39083
- HOTREQ-1070 Backport for SPARK-38318
- HOTREQ-1091 ITAU Casting invalid dates does not produce NULL
- HOTREQ-1114 Hue does not work with medium duty DL because IDBroker config has comma separated URLs

# Fixed Issues In Cloudera Runtime 7.2.14.5

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.5.

The following issues are resolved:

- HOTREQ-1178 - CDPD-44832 - HUE Oozie workflow rerun fails.
- HOTREQ-1161 - CFM - NIFI nodes are getting disconnected frequently

# Fixed Issues In Cloudera Runtime 7.2.14.6

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.6.

The following issues are resolved:

- HOTREQ-1224 - include CSA-3742 into CSA-DH
- HOTREQ-1201 - HADOOP-18476 - ABFS and S3A FileContext bindings to close wrapped filesystems in finalizer
- HOTREQ-1222 - ABFS - Disable readAhead for 7.2.12 and higher versions

## 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.

# Fixed Issues In Cloudera Runtime 7.2.14.7

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.7.

The following issues are resolved:

- HOTREQ-1264
  - CDPD-39592 - Due to the changes made in azure conf.py via PR-2396, Azure FB is unaccessible.
  - CDPD-29225 - Hue does not work with medium duty DL because IDBroker config has comma separated URLs

### CVE

- Upgrade Apache Commons Text to 1.10.0 due to CVE-2022-42889

### Known Issue OPSAPS-66017

Description: The datalake cluster goes to "Node failure" state after upgrading to the latest version. The HMS service goes to error state post upgrade.

Workaround: Restart the HMS service in datalake using the Cloudera Manager UI. The cluster health becomes normal after sometime.

# Fixed Issues In Cloudera Runtime 7.2.14.8

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.8.

The following issues are resolved:

- HOTREQ-1320 - HOTFIX for Bug - Add delegation token support for long running spark job
- HOTREQ-1321 - Hotfix for CVE-2022-25168 in CDP 7.2.11 version

### Known Issue CDPD-52789

**CDPD-52789: After performing the Datalake upgrade, HMS service goes to error state. The cluster event logs show the following error: Cloudera Manager reported health issues with node(s):...[The following services are in bad health: hive...]**

Using Cloudera Manager, restart the Hive MetaStore service. Wait for few minutes, the error clears automatically.

# Fixed Issues In Cloudera Runtime 7.2.14.9

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.14.9.

The following issue is resolved:

- CDPD-28710 - Backport HIVE-25380.

### Known Issue CDPD-52789

**CDPD-52789: After performing the Datalake upgrade, HMS service goes to error state. The cluster event logs show the following error: Cloudera Manager reported health issues with node(s):...[The following services are in bad health: hive...]**

Using Cloudera Manager, restart the Hive MetaStore service. Wait for few minutes, the error clears automatically.

# Service Pack in Cloudera Runtime 7.2.14

You can review the list of CDP Public Cloud hotfixes rolled into Cloudera Runtime 7.2.14. This will help you to verify if a hotfix provided to you on a previous CDP Public Cloud release was included in this release.

- HOTREQ-834 - Hotfix request - Avro
- HOTREQ-785 - Fix stackoverflow error in HiveMetaStore::get_partitions_by_names
- HOTREQ-917 - Hotfix Request for TSB 2022-543 for Azure Environment
- HOTREQ-889 - Multiple version of woodstox-core jars on the Spark classpath
- HOTREQ-888 - Hotfix for - IMPALA-11200

- HOTREQ-882 - Backport YARN-11020

**Note:** For more information about the updates related to Cloudera Manager 7.6.0, see Cloudera Manager Release Notes.

# Known Issues In Cloudera Runtime 7.2.14

You must be aware of the known issues and limitations, the areas of impact, and workaround in Cloudera Runtime 7.2.14.

## Known Issues in Apache Atlas

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

**DOCS-12597: NiFi service user's Atlas interaction (entity_read operations) filling Ranger audit collectionNone.**

Problem: NiFi frequently interacts with the Atlas API, which generates Audit logs of the API calls. These Audit logs are indexed in SOLR, which is backed by statically sized EBS volumes on the SDX core nodes (3*250 GB). The amount of these logs causes the EBS volumes to fill up, which takes the SOLR service down

Workaround: To reduce disk space consumption:

1. Reduce the SOLR TTL of the Ranger-audit collection
2. Increase the EBS volumes to their maximum of 1TB
3. Apply a cm_atlas audit filter to filter READ events from the NiFi service user against the Atlas API.

**DOCS-12697: Atlas canary check is disabled by default**

Problem: Check for Atlas was disabled by default, hence Cloudera Manager might not show proper health alerts.

Workaround: You must manually enable the canary check from Atlas configurations

**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-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

## Known Issues in Apache Avro

This topic describes known issues and workarounds for using Avro in this release of Cloudera Runtime.

**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 Cruise Control, the impact or changes to the functionality, and the workaround.

**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.

- 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 HBase, the impact or changes to the functionality, and the workaround.

**You cannot use the OpDB Data Hub template to create new clusters**

To resolve this issue, you can use COD or custom templates.

## Technical Service Bulletins

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

When using Replication Manager for Apache HBase (HBase) snapshot replication, data loss will occur if both of the following conditions are met: (i) the external account used for the operation has delete access to the target storage location, and (ii) the snapshot export fails. If these conditions are met, the cleanup operation, which is automatically performed after the failure, would delete all data in the root folder of the snapshot, not only the snapshot files. If the user account does not have the delete permission on the target folder, the data remains unaffected.

**Knowledge article**

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

# 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.

**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-13636: Hive job fails with OutOfMemory exception in the Azure DE cluster**

Set the parameter hive.optimize.sort.dynamic.partition.threshold=0. Add this parameter in Cloudera Manager (Hive Service Advanced Configuration Snippet (Safety Valve) for hive-site.xml)

**ENGESC-2214: Hiveserver2 and HMS service logs are not deleted**

Update Hive log4j configurations. Hive -> Configuration -> HiveServer2 Logging Advanced Configuration Snippet (Safety Valve) Hive Metastore -> Configuration -> Hive Metastore Server Logging Advanced Configuration Snippet (Safety Valve) Add the following to the configurations: appender.DRFA.strategy.action.type=DELETE appender.DRFA.strategy.action.basepath=${log.dir} appender.DRFA.strategy.action.maxdepth=1 appender.DRFA.strategy.action.PathConditions.glob=${log.file}.* appender.DRFA.strategy.action.PathConditions.type=IfFileName appender.DRFA.strategy.action.PathConditions.nestedConditions.type=IfAccumulatedFileCount appender.DRFA.strategy.action.PathConditions.nestedConditions.exceeds=same value as appender.DRFA.strategy.max

**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

**CDPD-11890: Hive on Tez cannot run certain queries on tables stored in encryption zones**

This problem occurs when the Hadoop Key Management Server (KMS) connection is SSL-encrypted and a self signed certificate is used. SSLHandshakeException might appear in Hive logs.

Use one of the workarounds:

- Install a self signed SSL certificate into cacerts file on all hosts.
- Copy ssl-client.xml to a directory that is available in all hosts. In Cloudera Manager, in Clusters Hive on Tez Configuration . In Hive Service Advanced Configuration Snippet for hive-site.xml, click +, and add the name tez.aux.uris and valuepath-to-ssl-client.xml.

## Technical Service Bulletins

**TSB 2022-567: Potential Data Loss due to CTLT HBaseStorageHandler failure dropping underlying HBase table while rollback**

If the create table target_table like source table command (CTLT) fails and the source table is HBaseStorageHandler-based table, the HBaseMetaHook rollback logic deletes the underlying HBase table, resulting in potential data loss.

**Upstream JIRA**

HIVE-25989

**Knowledge article**

For the latest update on this issue, see the corresponding Knowledge article: TSB 2022-567: Potential Data Loss due to CTLT HBaseStorageHandler failure dropping underlying HBase table while rollback

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

An IN or an OR predicate involving a binary datatype column may produce wrong results. The OR predicate is converted to an IN due to the setting hive.optimize.point.lookup which is true by default. Only binary data types are affected by this issue. See https://issues.apache.org/jira/browse/HIVE-26235 for example queries which may be affected.

**Upstream JIRA**

HIVE-26235

**Knowledge article**

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

# Known Issues in Hue

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

**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***]
```

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

In CDP, Hue is compatible with Python 2.7.x, but the Tablib library for Hue has been upgraded from 0.10.x to 0.14.x, which is generally used with the Python 3 release. If you try to download Impala query results having special characters in the result set in a CSV format, then the download may fail with the ASCII unicode decode error.

To fix this issue, downgrade the Tablib library to 0.12.x.

1. SSH into the Hue server host.
2. Change directory to the following:

```
cd /opt/cloudera/parcels/CDH-7.x/lib/
```

3. Back up the hue directory:

```
cp -R hue hue_orginal
```

4. Change to the hue directory:

```
cd hue
```

5. Install the Wheel package using pip:

```
./build/env/bin/pip install wheel
```

The Wheel package is used to avoid recompiling your software during every install.

6. Install the Python Setuptools package for Hue as follows:

```
./build/env/bin/pip install setuptools==44.1.0
```

7. Install Tablib version 0.12.1 as follows:

```
./build/env/bin/pip install tablib==0.12.1
```

8. Go to Cloudera Manager and restart the Hue service.

### Impala SELECT table query fails with UTF-8 codec error

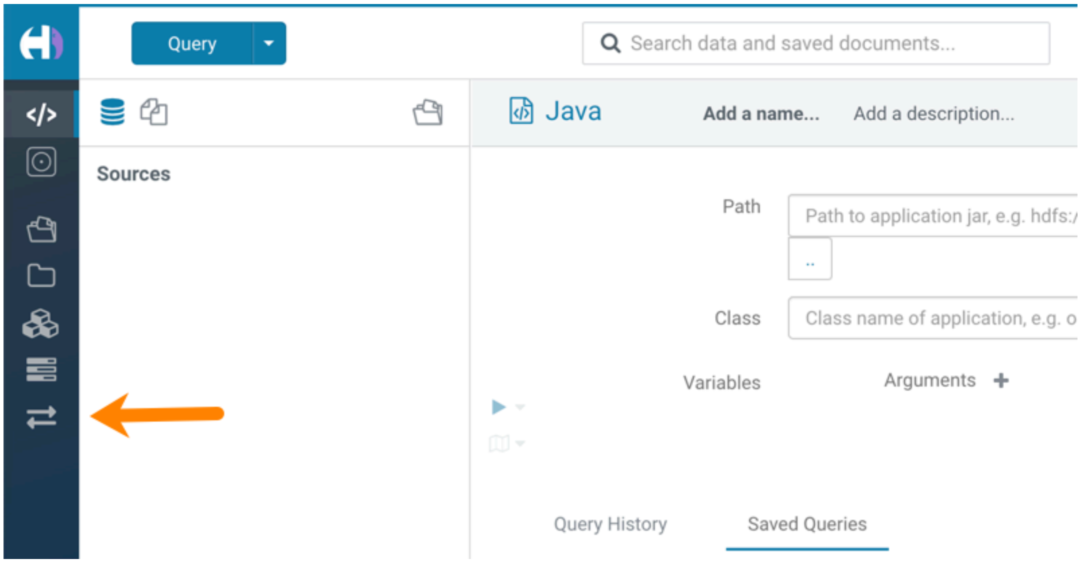
Hue cannot handle columns containing non-UTF8 data. As a result, you may see the following error while queying tables from the Impala editor in Hue: 'utf8' codec can't decode byte 0x91 in position 6: invalid start byte.

To resolve this issue, contact Cloudera Support to apply the following software patch: ENGESC-3457.

### 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



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

You may see the following error message while starting the Hue Load Balancer:

```
BalancerMember worker hostname (xxx-xxxxxxxx-xxxxxxxxxxx-xxxxxxx
.xxxxxx-xxxxxx-xxxxxx.example.site) too long.
```

Or, the Hue load balancer role fails to start after the upgrade, which prevents the Hue service from starting. If this failure occurs during cluster creation, cluster creation fails with the following error:

```
com.sequenceiq.cloudbreak.cm.ClouderaManagerOperationFailedExcep
tion: Cluster template install failed: [Command [Start], with id
 [1234567890] failed:
Failed to start role., Command [Start], with id [1234567890] fail
ed: Failed to start role., Command [Start], with id [1234567890]
 failed: Failed to start role.]
Unable to generate configuration for HUE_SERVER
Role failed to start due to error com.cloudera.cmf.service.confi
g.ConfigGenException: Unable to generate config file hue.ini
```

Cloudera Manager displays this error when you create a Data Hub cluster using the Data Engineering template and the Hue Load Balancer worker node name has exceeded 64 characters. In a CDP Public Cloud deployment, the system automatically generates the Load Balancer worker node name through AWS or Azure.

For example, if you specify cdp-123456-scalecluster as the cluster name, CDP creates cdp-123456-scalecluster-master2.repro-aw.a123-4a5b.example.site as the worker node name.

Specify a shorter cluster name while creating a Data Hub cluster so that the final worker node name does not cross 64 characters.

For example, cdp-123456-scale.

### Unsupported features

#### 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 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.

## Technical Service Bulletins

### TSB 2021-479: Impala can return incomplete results through JDBC and ODBC clients in all CDP offerings

In CDP, we introduced a timeout on queries to Impala defaulting to 10 seconds. The timeout setting is called FETCH_ROWS_TIMEOUT_MS. Due to this setting, JDBC, ODBC, and Beeswax clients running Impala queries believe the data returned at 10 seconds is a complete dataset and present it as the final output. However, in cases where there are still results to return after this timeout has passed, when the driver closes the connection, based on the timeout, it results in a scenario where the query results are incomplete.

### Upstream JIRA

IMPALA-7561

### Knowledge article

For the latest update on this issue, see the corresponding Knowledge article: TSB-2021 479: Impala can return incomplete results through JDBC and ODBC clients in all CDP offerings

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

Apache Impala may produce incorrect results for a query which has all of the following conditions:

- There are two or more analytic functions (for example, row_number()) in an inline view
- Some of the functions have partition-by expression while the others do not
- There is a predicate on the inline view's output expression corresponding to the analytic function

**Upstream JIRA**

IMPALA-11030

**Knowledge article**

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

# 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.

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

By default all created topics are secured. However, when topic creation and deletion is done with the kafka-topics tool using the --zookeeper option, the tool talks directly to Zookeeper. Because security is the responsibility of ZooKeeper authorization and authentication, Kafka cannot prevent users from making ZooKeeper changes. As a result, if the --zookeeper option is used, only the user who created the topic will be able to carry out administrative actions on it. In this scenario Kafka will not have permissions to perform tasks on topics created this way.

Use kafka-topics with the --bootstrap-server option that does not require direct access to Zookeeper.

**Certain Kafka command line tools require direct access to Zookeeper**

The following command line tools talk directly to ZooKeeper and therefore are not secured via Kafka:

- kafka-reassign-partitions

None

**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-62209: Kafka rolling restart checks fail when non-default hostnames are used**

When Cloudera Manager runs Kafka broker rolling restart checks, it uses kafka-topics commands to gather information about the brokers. When running the commands, Cloudera Manager uses the default bootstrap server (host:port pair) that is automatically configured by Cloudera Manager. If your Kafka brokers are configured to use custom listeners and you use non-default hosts in the listener configuration, Cloudera Manager will be unable to establish a connection with the brokers. As a result, both the rolling restart check and the rolling restart fails.

None

**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

## 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.
- While Kafka Connect is available as part of Runtime, it is considered technical preview and is currently not supported in CDP Public Cloud. NiFi is a proven solution for batch and real time data loading that complement Kafka's message broker capability. For more information, see Creating your first Flow Management cluster.
- 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.

## 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

## Known Issues in Apache Oozie

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

**CDPD-29302: The Atlas lineage information is missing in case of HWC JDBC write.**

None

**CDPD-29297: HWC + Oozie issue: Cannot create PoolableConnectionFactory**

Currently only Spark cluster mode is supported in the Oozie Spark Action with Hive Warehouse Connector (HWC).

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.

### CDPD-5340: The resourceManager property defined in an Oozie workflow might not work properly if the workflow is submitted through Knox proxy.

An Oozie workflow defined to use the resourceManager property might not work as expected in situations when the workflow is submitted through Knox proxy.

Define the jobTracker property with the same value as that of the resourceManager property.

### 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 7.2.14.

## 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.

**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.

**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 2023-644: Microsoft Azure parent directory deletion**

Cloudera has observed that, in CDP Public Cloud environments using ADLS file system, it is possible to delete the parent folder even though the Ranger permission is given only for the subfolder. If a user has delete permissions on a particular folder, this user can also delete the parent folder containing the subfolder for which the user has the delete permission.

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

## Known Issues in Schema Registry

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

**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-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-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.

## Known Issues

**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.

**CDPD-24003: The Solr admin UI is only accessible with full solr_admin permission in Ranger**

Full solr_admin permission is required in Ranger to access the Solr Admin UI.

None.

**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.

If the timeout occurs during Near real time (NRT) indexing, Cloudera suggests you 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 solrClientSocketTimeout 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-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 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
```

**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.

**Solr service with no added collections causes the upgrade process to fail**

Upgrade fails while performing the bootstrap collections step of the solr-upgrade.sh script with the error message:

```
Failed to execute command Bootstrap Solr Collections on service
Solr
```

if there are no collections present in Solr.

If there are no collections added to it, remove the Solr service from your cluster before you start the upgrade.

**CDH-34050: Collection Creation No Longer Supports Automatically Selecting A Configuration If Only One Exists**

Before CDH 5.5.0, a collection could be created without specifying a configuration. If no -c value was specified, then:

- If there was only one configuration, that configuration was chosen.
- If the collection name matched a configuration name, that configuration was chosen.

Search now includes multiple built-in configurations. As a result, there is no longer a case in which only one configuration can be chosen by default.

Explicitly specify the collection configuration to use by passing -c <configName> to solrctl coll ection --create.

**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-19923: The quickstart.sh file does not validate ZooKeeper and the NameNode on some operating systems.**

The quickstart.sh file uses the timeout function to determine if ZooKeeper and the NameNode are available. To ensure this check can be complete as intended, the quickstart.sh determines if the operating system on which the script is running supports timeout. If the script detects that the operating system does not support timeout, the script continues without checking if the NameNode and ZooKeeper are available. If your environment is configured properly or you are using an operating system that supports timeout, this issue does not apply.

This issue only occurs in some operating systems. If timeout is not available, the quickstart continues and final validation is always done by the MapReduce jobs and Solr commands that are run by the quickstart.

**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.

**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 the Cloudera Search product 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 workaround.

**CDPD-217: HBase/Spark connectors are not supported**

The *Apache HBase Spark Connector* (hbase-connectors/spark) and the *Apache Spark - Apache HBase Connector* (shc) are not supported in the initial CDP release.

None

**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.

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

Use a different compression algorithm.

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

Apache Ranger in CDP disallows Spark users from running CREATE VIEW statements.

Create the view using Hive or Impala.

**CDPD-3783: Cannot create databases from Spark**

Attempting to create a database using Spark results in an error similar to the following:

```
org.apache.spark.sql.AnalysisException:
            org.apache.hadoop.hive.ql.metadata.HiveException: Me
taException(message:Permission denied: user [sparkuser] does not
 have [ALL] privilege on [hdfs://ip-10-1-2-3.cloudera.site:8020/
tmp/spark/warehouse/spark_database.db]);
```

Create the database using Hive or Impala, or specify the external data warehouse location in the create command. For example:

```
sql("create database spark_database location '/warehouse/tablesp
ace/external/hive/spark_database.db'")
```

## Known Issues for Apache Sqoop

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

**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 7.2.14.

**CDPD-33770: On the topics details page selecting a custom timestamp is broken**

When you select a custom TimePeriod (a non-predefined TimePeriod like 6 hours, 30 minutes etc.) on the SMM UI's topicDetail page, an error is going to be thrown, and the replication related metrics would not be displayed.

**OPSAPS-63017: The Kafka Connect tab is missing from the SMM UI**

Under certain circumstances the Kafka Connect tab in SMM might not be available by default on Data Hub clusters even if Kafka Connect is provisioned on the cluster. As a result, interacting with Kafka Connect using SMM is not possible.

1. Access the Cloudera Manager instance managing the affected Data Hub cluster.
2. Select the Streams Messaging Manager service, and go to Configuration.
3. Find and configure the following properties:
   - Kafka Connect Host

     Enter the hostname of the machine that the Kafka Connect role is deployed on. If you have multiple instances of the Kafka Connect role, you can choose to use any of them. Add a single hostname, as configuring multiple hostnames for high availability is currently not supported.
   - Kafka Connect Port

     Enter the port that the Kafka Connect role is using. The value of this property must match the port set in the Secure Kafka Connect Rest Port Kafka property.
   - Kafka Connect Protocol

     Set this property to https.

**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.

**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.

# 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-63104: The automatically generated password for co-located services is invalid**

SRM automatically generates a username and password that can be used by co-located services to access SRM and its REST API. However, a unique password is generated for each SRM Service role instance. Because of this, co-located services that use the password, for example SMM, can only connect to one of the SRM Service role instances.

Manually configure a password using the SRM Service Co-Located Service User Password SRM property. The password you configure will be accepted by all SRM Service role instances.

**OPSAPS-63992: Rolling restart unavailable for SRM**

Initiating a rolling restart for the SRM service is not possible. Consequently, performing a rolling upgrade of the SRM service is also not possible.

None

**CDPD-31745: SRM Control fails to configure internal topic when target is earlier than Kafka 2.3**

When the target Kafka cluster of a replication is earlier than version 2.3, the srm-control internal topic is created with an incorrect configuration (cleanup.policy=compact). This causes the srm-control topic to lose the replication filter records, causing issues in the replication.

After a replication is enabled where the target Kafka cluster is earlier than 2.3, manually configure all srm-control.*[***SOURCE CLUSTER ALIAS***]*.internal topics in the target cluster to use cleanup.policy=compact.

### 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 noexec 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-62546: Kafka External Account SSL keypassword configuration is used incorrectly by SRM

When a Kafka External Account specifies a keystore that uses an SSL key password, SRM uses it as the ssl.keystore.key configuration. Due to using the incorrect ssl.keystore.key configuration, SRM will fail to load the keystore in certain cases.

Workaround: For the keystores used by the Kafka External Accounts, the SSL key password should match the SSL keystore password, and the SSL keystore key password should not be provided. Alternatively, you can use the legacy connection configurations based on the streams.replication.manager.configs to specify the SSL key password.

## 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.

### SRM cannot ensure the exactly-once semantics of transactional source topics

SRM data replication uses at-least-once guarantees, and as a result cannot ensure the exactly-once semantics (EOS) of transactional topics in the backup/target cluster.

**Note:** Even though EOS is not guaranteed, you can still replicate the data of a transactional source, but you must set isolation.level to read_committed for SRM's internal consumers. This can be done by adding *[***CONFIG LEVEL PREFIX***]*.isolation.level=read_committed to the Streams Replication Manager's Replication Configs SRM service property in Cloudera Manger. The isolation.level property can be set on a global connector or replication level. For example:

```
#Global connector level
connectors.consumer.isolation.level=read_committed
#Replication level
uswest->useast.consumer.isolation.level=read_committed
```

**SRM checkpointing is not supported for transactional source topics**

SRM does not correctly translate checkpoints (committed consumer group offsets) for transactional topics. Checkpointing assumes that the offset mapping function is always increasing, but with transactional source topics this is violated. Transactional topics have control messages in them, which take up an offset in the log, but they are never returned on the consumer API. This causes the mappings to decrease, causing issues in the checkpointing feature. As a result of this limitation, consumer failover operations for transactional topics is not possible.

## Known Issues in MapReduce and YARN

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

### Known Issues

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

The Queue Manager UI may show incorrect value for the Maximum Parallel Applications property for the root queue. The value can be changed with the QM UI, but the UI will always show 2147483647.

Check the Maximum Parallel Applications property for the root queue manually:

1. In Cloudera Manager, select to the YARN service.
2. Click Configuration.
3. Find the Capacity Scheduler Configuration Advanced Configuration Snippet (Safety Valve) property.
4. Find the yarn.scheduler.capacity.root.max-parallel-apps property in the safety valve.

**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.

**COMPX-5244: Root queue should not be enabled for auto-queue creation**

After dynamic auto child creation is enabled for a queue using the YARN Queue Manager UI, you cannot disable it using the YARN Queue Manager UI. That can cause problem when you want to switch between resource allocation modes, for example from weight mode to relative mode. The YARN Queue Manager UI does not let you to switch resource allocation mode if there is at least one dynamic auto child creation enabled parent queue in your queue hierarchy.

If the dynamic auto child creation enabled parent queue is NOT the root or the root.default queue: Stop and remove the dynamic auto child creation enabled parent queue. Note that this stops and remove all of its child queues as well.

If the dynamic auto child creation enabled parent queue is the root or the root.default queue: You cannot stop and remove neither the root nor the root.default queue. You have to change the configuration in the applicable configuration file:

1. In Cloudera Manager, navigate to YARN>>Configuration.
2. Search for capacity scheduler and find the Capacity Scheduler Configuration Advanced Configuration Snippet (Safety Valve) property.
3. Add the following configuration: yarn.scheduler.capacity.<queue-path>.auto-queue-creation-v2.enabled=false For example: yarn.scheduler.capacity.root.default.auto-queue-creation-v2.enabled=false Alternatively, you can remove the yarn.scheduler.capacity.<queue-path>.auto-queue-creation-v2.enabled property from the configuration file.

4. Restart the Resource Manager.

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

Scenario

1. User creates one or more partitions.
2. Assigns a partition to a parent with children
3. Switches to the partition to distribute the capacities
4. Creates a new child queue under one of the leaf queues but the following error is displayed:

```
Error :
2021-03-05 17:21:26,734 ERROR
com.cloudera.cpx.server.api.repositories.SchedulerRepository: Val
idation failed for Add queue
operation. Error message: CapacityScheduler configuration vali
dation failed:java.io.IOException:
Failed to re-init queues : Parent queue 'root.test2' have childr
en queue used mixed of  weight
mode, percentage and absolute mode, it is not allowed, please do
uble check, details:
{Queue=root.test2.test2childNew, label= uses weight mode}. {Que
ue=root.test2.test2childNew,
label=partition uses percentage mode}
```

To create new queues under leaf queues without hitting this error, perform the following:

1. Switch to Relative mode
2. Create the required queues
3. Create the required partitions
4. Assign partitions and set capacities
5. Switch back to Weight mode

1. Create the entire queue structure
2. Create the required partitions
3. Assign partition to queues
4. Set partition capacities

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

In the current implemention, if there is an existing managed queue in Relative mode, then conversion to Weight mode is not be allowed.

To proceed with the conversion from Relative mode to Weight mode, there should not be any managed queues. You must first delete the managed queues before conversion. In Weight mode, a parent queue can be converted into managed parent queue.

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

If you associate a partition with one or more queues and then switch the allocation mode before assigning capacities to the queues, an Operation Failed error is displayed as the `max-capacity` is set to null.

After you associate a partition with one or more queues, in the YARN Queue Manager UI, click Overview > <*PARTITION NAME*> from the dropdown list and distribute capacity to the queues before switching allocation mode or creating placement rules.

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

If you delete a partition (node label) which has been associated with queues and those queues have capacities configured for that partition (node label), the CS.xml still contains the partition (node label) information. Hence, you cannot switch to absolute mode after deleting the partition (node label).

It is recommended not to delete a partition (node label) which has been associated with queues and those queues have capacities configured for that partition (node label).

**COMPX-3181: Application logs does not work for AZURE and AWS cluster**

Yarn Application Log Aggregation will fail for any YARN job (MR, Tez, Spark, etc) which do not use cloud storage, or use a cloud storage location other than the one configured for YARN logs (`yarn.nodemanager.remote-app-log-dir`).

Configure the following:

- For MapReduce job, set mapreduce.job.hdfs-servers in the mapred-site.xml file with all filesystems required for the job including the one set in yarn.nodemanager.remote-app-log-dir such as hdfs://nn1/,hdfs://nn2/.
- For Spark job, set the job level with all filesystems required for the job including the one set in yarn.nodemanager.remote-app-log-dir such as hdfs://nn1/,hdfs://nn2/ in spark.yarn.access.hadoopFileSystems and pass it through the `--config` option in `spark-submit`.
- For jobs submitted using the hadoop command, place a separate core-site.xml file with fs.defaultFS set to the filesystem set in yarn.nodemanager.remote-app-log-dir in a path. Add that directory path in `--config` when executing the hadoop command.

**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-1451: Queue Manager does not support multiple ResourceManagers**

When YARN High Availability is enabled there are multiple ResourceManagers. Queue Manager receives multiple ResourceManager URLs for a High Availability cluster. It picks the active ResourceManager URL only when Queue Manager page is loaded. Queue Manager cannot handle it gracefully when the currently active ResourceManager goes down while the user is still using the Queue Manager UI.

Reload the Queue Manager page manually.

**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.

**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.

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

This issue does not affect the functionality. Restarting Yarn service will fix this issue.

### 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-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.

### CDH-6808: Routable IP address required by ResourceManager

ResourceManager requires routable host:port addresses for yarn.resourcemanager.scheduler.addre ss, 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.

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

Stacks under the Logs directory for Hadoop daemons, such as NameNode, DataNode, ResourceManager, NodeManager, and JobHistoryServer are not accessible from Knox Gateway.

Administrators can SSH directly to the Hadoop Daemon machine to collect stacks under the Logs directory.

### 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

### OPSAPS-50291: Environment variables HADOOP_HOME, PATH, LANG, and TZ are not getting whitelisted

It is possible to include the environment variables HADOOP_HOME, PATH, LANG, and TZ in the allowlist, but the container launch environments do not have these variables set up automatically.

You can manually add the required environment variables to the allowlist using Cloudera Manager.

1. In Cloudera Manager, select the YARN service.
2. Click the Configuration tab.
3. Search for Containers Environment Variable Whitelist.
4. Add the environment variables (HADOOP_HOME, PATH, LANG, TZ) which are required to the list.
5. Click Save Changes.
6. Restart all NodeManagers.
7. Check the YARN aggregated logs to ensure that newly whitelisted environment variables are set up for container launch.

### 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.

**COMPX-8687: Missing access check for getAppAttemps**

When the Job ACL feature is enabled using Cloudera Manager ( YARN Configuration Enablg JOB ACL property), the mapreduce.cluster.acls.enabled property is not generated to all configuration files, including the yarn-site.xml configuration file. As a result the ResourceManager process will use the default value of this property. The default property of mapreduce.cluster.acls.enabled is false.

Workaround: Enable the Job ACL feature using an advanced configuration snippet:

1. In Cloudera Manager select the YARN service.
2. Click Configuration.
3. Find the YARN Service MapReduce Advanced Configuration Snippet (Safety Valve) property.
4. Click the plus icon and add the following:
   - Name: mapreduce.cluster.acls.enabled
   - Value: true
5. Click Save Changes.

### 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
- Moving jobs between queues
- 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

## Known Issues in Apache Zeppelin

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

**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 7.2.14

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 7.2.14.

## 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 default TLS protocol when using Java 11 or later is updated to TLS 1.3. TLS 1.2 remains the default for earlier Java versions. For more information, see KIP-573.

Previous behavior:

The default TLS protocol used was TLS 1.2.

New behavior:

The default TLS protocol when using Java 11 or later is TLS 1.3.

**Summary:**

The client.dns.lookup client property is now set to use_all_dns_ips by default. With this option set, clients attempt to connect to the broker using all possible IP addresses of a hostname. This option reduces connection failure rates and is better suited for cloud and containerized environments where a single hostname might resolve to multiple IP addresses. For more information, see KIP-602.

Previous behavior:

The default value of the client.dns.lookup client property was set to default.

New behavior:

The default value of the client.dns.lookup property is set to use_all_dns_ips.

**Summary:**

The default value for the following Kafka Connect properties is increased to 3:

- Offset Storage Topic Replication Factor(offset.storage.replication.factor)
- Configuration Storage Topic Replication Factor(config.storage.replication.factor))
- Status Storage Topic Replication Factor (status.storage.replication.factor)

Previous behavior:

The default value of the properties was set to 1.

New behavior:

The default value of the properties is set to 3.

## Behavioral Changes in Cloudera Search

Learn about the change in certain functionality of Cloudera Search that has resulted in a change in behavior from the previously released version to this version of Cloudera Runtime.

**Summary:**

Validation of javax.security.auth.useSubjectCredsOnly parameter changed in Solr client

Previous behavior:

Unless explicitely set to 'true', Solr client used to set the value of javax.security.auth.useSubjectCredsOnly parameter to 'false'.

New behavior:

Solr client does not overwrite ‘null’ value of javax.security.auth.useSubjectCredsOnly parameter, it only throws a warning during connection setup:

```
System Property: javax.security.auth.useSubjectCredsOnly set to:
 [true|null] not false.
SPNego authentication may not be successful.
```

This may cause issues when connecting Solr to custom applications. To prevent this, set

```
-Djavax.security.auth.useSubjectCredsOnly=false
```

in the JVM configuration of those applications.

**Note:** For the Spark-Solr connector, use

```
spark-submit --driver-java-options "-Djavax.security
.auth.useSubjecCredsOnly=false"
```

Cloudera has implemented this change in MapReduceIndexer, CrunchIndexer, Hive-Solr connector, and Atlas.

# Deprecation Notices In Cloudera Runtime 7.2.14

Certain features and functionalities have been removed or deprecated in Cloudera Runtime 7.2.14. 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 in Apache Kudu

Certain features and functionality in Kudu are deprecated or removed in Cloudera Runtime 7.2.14. You must review these changes along with the information about the features in Kudu that will be removed or deprecated in a future release.

- The Flume sink has been migrated to the Apache Flume project and removed from Kudu. Users depending on the Flume integration can use the old kudu-flume jars or migrate to the Flume jars containing the Kudu sink.
- Support for Apache Sentry authorization has been deprecated and may be removed in the next release. Users depending on the Sentry integration should migrate to the Apache Ranger integration for authorization.
- Support for Python 2 has been deprecated and may be removed in the next release.
- Support for CentOS/RHEL 6, Debian 8, Ubuntu 14 has been deprecated and may be removed in the next release.

## Deprecation Notices for Apache Kafka

Certain features and functionality in Apache Kafka are deprecated or removed in Cloudera Runtime 7.2.14. You must review these changes along with the information about the features in Kafka that will be removed or deprecated in a future release.

### Deprecated

**kafka-preferred-replica-election**

The kafka-preferred-replica-election.sh command line tool has been deprecated in upstream Apache Kafka 2.4.0. Its alternative in CDP, kafka-preferred.replica-election, is also deprecated.

**--zookeeper**

The --zookeeper option has been deprecated for all Kafka command line tools except for kafka-re assign-partitions. Cloudera recommends that you use the --bootstrap-server option instead.