

Cloudera Runtime 7.1.4

Release Notes

Date published: 2020-10-13

Date modified:

CLOUdera

<https://docs.cloudera.com/>

Legal Notice

© Cloudera Inc. 2025. All rights reserved.

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

Overview.....	5
Cloudera Runtime Component Versions.....	5
Using the Cloudera Runtime Maven Repository.....	6
Maven Artifacts for Cloudera Runtime 7.1.4.0.....	7
What's New In Cloudera Runtime 7.1.4.....	25
What's New in Apache Atlas.....	25
What's New in Cruise Control.....	25
What's new in Data Analytics Studio.....	25
What's New in Apache HBase.....	25
What's New in Apache Hadoop HDFS.....	25
What's New in Apache Hive.....	25
What's New in Hue.....	25
What's New in Apache Impala.....	26
What's New in Apache Kafka.....	26
What's New in Apache Knox.....	26
What's New in Apache Kudu.....	26
What's New in Apache Oozie.....	26
What's New in Apache Hadoop Ozone.....	27
What's New in Apache Phoenix.....	27
What's New in Schema Registry.....	27
What's New in Cloudera Search.....	27
What's New in Sqoop.....	27
What's New in Streams Replication Manager.....	27
What's new in Streams Messaging Manager.....	28
What's New in Apache Hadoop YARN.....	28
What's New in Apache ZooKeeper.....	28
Fixed Issues In Cloudera Runtime 7.1.4.....	28
Fixed Issues in Apache Atlas.....	28
Fixed issues in Data Analytics Studio.....	29
Fixed Issues in Apache Hadoop.....	29
Fixed Issues in HDFS.....	29
Fixed Issues in Apache HBase.....	30
Fixed Issues in Apache Hive.....	30
Fixed Issues in Hue.....	31
Fixed Issues in Apache Impala.....	31
Fixed Issues in Apache Kafka.....	32
Fixed Issues in Apache Kudu.....	32
Fixed Issues in Apache Oozie.....	32
Fixed issues in Ozone.....	33
Fixed Issues in Phoenix.....	34
Fixed Issues in Apache Ranger.....	34

Fixed Issues in Schema Registry.....	35
Fixed Issues in Cloudera Search.....	35
Fixed Issues in Spark.....	35
Fixed Issues in Apache Sqoop.....	36
Fixed Issues in Streams Replication Manager.....	36
Fixed Issues in Streams Messaging Manager.....	36
Fixed Issues in Apache YARN.....	38
Fixed Issues in Zeppelin.....	38

Hotfixes in Cloudera Runtime 7.1.4..... 38

Known Issues In Cloudera Runtime 7.1.4..... 38

Known Issues in Apache Atlas.....	38
Known issues in Cruise Control.....	42
Known Issues in Data Analytics Studio.....	42
Known Issues in Apache Hadoop.....	44
Known Issues in Apache HBase.....	44
Known Issues in HDFS.....	46
Known Issues in Apache Hive.....	47
Known Issues in Hue.....	51
Known Issues in Apache Impala.....	55
Known Issues in Apache Kafka.....	61
Known Issues in Kerberos.....	63
Known Issues in Apache Knox.....	63
Known Issues in Apache Kudu.....	64
Known Issues in Apache Oozie.....	64
Known Issues in Ozone.....	65
Known Issues in Apache Phoenix.....	66
Known Issues in Apache Ranger.....	66
Known Issues in Schema Registry.....	67
Known Issues in Cloudera Search.....	67
Known Issues in Apache Solr.....	70
Known Issues in Apache Spark.....	70
Known Issues in Streams Replication Manager.....	71
Known Issues for Apache Sqoop.....	73
Known issues in Streams Messaging Manager.....	74
Known Issues in MapReduce and YARN.....	75
Known Issues in Apache Zeppelin.....	80
Known Issues in Apache ZooKeeper.....	80

Behavioral Changes In Cloudera Runtime 7.1.4.....81

Behavioral Changes in Apache Kudu.....	81
Behavioral Changes in Apache Phoenix.....	81
Behavioral Changes in Ranger.....	81
Behavioral Changes in Cloudera Search.....	82

Deprecation Notices In Cloudera Runtime 7.1.4.....82

Deprecation notices in Apache Kudu.....	83
Deprecation Notices for Apache Kafka.....	83
Deprecation Notices in Apache HBase.....	83


Overview

You can review the Release Notes of Cloudera Runtime 7.1.4 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 7.1.4 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	0.8.0
Apache Atlas	2.1.0
Apache Calcite	1.19.0
Apache Avro	1.8.2
Apache Hadoop (Includes YARN and HDFS)	3.1.1
Apache HBase	2.2.3
Apache Hive	3.1.3000
Apache Impala	3.4.0
Apache Kafka	2.4.1
Apache Knox	1.3.0
Apache Kudu	1.12.0
Apache Livy	0.6.0
Apache MapReduce	3.1.1
Apache Ozone	0.5.0
Apache Oozie	5.1.0
Apache ORC	1.5.1
Apache Parquet	1.10.99
Apache Phoenix	5.0.0
Apache Ranger	2.0.0
Apache Solr	8.4.1
Apache Spark	<div>  Note: The version string reported by the software in this release is incorrect. Although the Apache Spark component of the version string indicates that it is based on Spark 2.4.0, the Spark component in Cloudera Runtime 7.1.4 is based on Apache Spark 2.4.5, not 2.4.0. </div> 2.4.5
Apache Sqoop	1.4.7
Apache Tez	0.9.1
Apache Zeppelin	0.8.2

Component	Version
Apache ZooKeeper	3.5.5

Other Components

Component	Version
Cruise Control	2.0.100
Data Analytics Studio	1.4.2
GCS Connector	1.9.10
HBase Indexer	1.5.0
Hue	4.5.0
Search	1.0.0
Schema Registry	0.8.1
Streams Messaging Manager	2.1.0
Streams Replication Manager	1.0.0

Connectors and Encryption Components

Component	Version
HBase connectors	1.0.0
Hive Meta Store (HMS)	1.0.0
Hive on Tez	1.0.0
Hive Warehouse Connector	1.0.0
Spark Atlas Connector	0.1.0
Spark Schema Registry	1.1.0

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 repository.cloudera.com.



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 7.1.4.0

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	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-aws-s3-bridge	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-classification-updater	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-client-common	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-client-v1	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-client-v2	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-common	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-distro	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-docs	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-graphdb-api	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-graphdb-common	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-graphdb-janus	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-index-repair-tool	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-intg	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-janusgraph-hbase2	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-notification	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-plugin-classloader	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-repository	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-server-api	2.0.0.7.1.4.0-203
	org.apache.atlas	atlas-testtools	2.0.0.7.1.4.0-203
	org.apache.atlas	hbase-bridge	2.0.0.7.1.4.0-203
	org.apache.atlas	hbase-bridge-shim	2.0.0.7.1.4.0-203
	org.apache.atlas	hbase-testing-util	2.0.0.7.1.4.0-203
	org.apache.atlas	hdfs-model	2.0.0.7.1.4.0-203
	org.apache.atlas	hive-bridge	2.0.0.7.1.4.0-203
	org.apache.atlas	hive-bridge-shim	2.0.0.7.1.4.0-203
	org.apache.atlas	impala-bridge	2.0.0.7.1.4.0-203
	org.apache.atlas	impala-bridge-shim	2.0.0.7.1.4.0-203
	org.apache.atlas	impala-hook-api	2.0.0.7.1.4.0-203
	org.apache.atlas	kafka-bridge	2.0.0.7.1.4.0-203
	org.apache.atlas	navigator-to-atlas	2.0.0.7.1.4.0-203
	org.apache.atlas	sqoop-bridge	2.0.0.7.1.4.0-203
	org.apache.atlas	sqoop-bridge-shim	2.0.0.7.1.4.0-203
Apache Avro	org.apache.avro	avro	1.8.2.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.avro	avro-compiler	1.8.2.7.1.4.0-203
	org.apache.avro	avro-ipc	1.8.2.7.1.4.0-203
	org.apache.avro	avro-mapred	1.8.2.7.1.4.0-203
	org.apache.avro	avro-maven-plugin	1.8.2.7.1.4.0-203
	org.apache.avro	avro-protobuf	1.8.2.7.1.4.0-203
	org.apache.avro	avro-service-archetype	1.8.2.7.1.4.0-203
	org.apache.avro	avro-thrift	1.8.2.7.1.4.0-203
	org.apache.avro	avro-tools	1.8.2.7.1.4.0-203
	org.apache.avro	trevni-avro	1.8.2.7.1.4.0-203
	org.apache.avro	trevni-core	1.8.2.7.1.4.0-203
Apache Calcite	org.apache.calcite	calcite-babel	1.19.0.7.1.4.0-203
	org.apache.calcite	calcite-core	1.19.0.7.1.4.0-203
	org.apache.calcite	calcite-druid	1.19.0.7.1.4.0-203
	org.apache.calcite	calcite-linq4j	1.19.0.7.1.4.0-203
	org.apache.calcite	calcite-server	1.19.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica	1.16.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica-core	1.16.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica-metrics	1.16.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica-metrics-dropwizardmetrics	1.16.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica-noop-driver	1.16.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica-server	1.16.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica-standalone-server	1.16.0.7.1.4.0-203
	org.apache.calcite.avatica	avatica-tck	1.16.0.7.1.4.0-203
Apache Crunch	org.apache.crunch	crunch-archetype	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-contrib	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-core	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-examples	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-hbase	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-hive	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-scrunch	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-spark	0.11.0.7.1.4.0-203
	org.apache.crunch	crunch-test	0.11.0.7.1.4.0-203
Apache Druid	org.apache.druid	druid-aws-common	0.17.1.7.1.4.0-203
	org.apache.druid	druid-benchmarks	0.17.1.7.1.4.0-203
	org.apache.druid	druid-console	0.17.1.7.1.4.0-203
	org.apache.druid	druid-core	0.17.1.7.1.4.0-203
	org.apache.druid	druid-gcp-common	0.17.1.7.1.4.0-203
	org.apache.druid	druid-hll	0.17.1.7.1.4.0-203
	org.apache.druid	druid-indexing-hadoop	0.17.1.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.druid	druid-indexing-service	0.17.1.7.1.4.0-203
	org.apache.druid	druid-integration-tests	0.17.1.7.1.4.0-203
	org.apache.druid	druid-processing	0.17.1.7.1.4.0-203
	org.apache.druid	druid-server	0.17.1.7.1.4.0-203
	org.apache.druid	druid-services	0.17.1.7.1.4.0-203
	org.apache.druid	druid-sql	0.17.1.7.1.4.0-203
	org.apache.druid	extendedset	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hadoop-vro-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-basic-security	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-bloom-filter	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-datasketches	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-ec2-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-google-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hdfs-storage	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-histogram	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-kafka-extraction-namespace	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-kafka-indexing-service	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-kerberos	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-kinesis-indexing-service	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-lookups-cached-global	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-lookups-cached-single	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-morc-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-parquet-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-protobuf-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-s3-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-stats	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-sql-metadata-storage	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-sql-metadata-storage	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-ssl-client-sslcontext	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-stdout-emitter	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-s3-stdout-emitter	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-s3-core-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-s3-aws-auth-storage	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-s3-local-files-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-slack-notification	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-splunkd-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-splunkdb-emitter	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-summarization-sketch	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-tomcat-http-average-query	0.17.1.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.druid.extensions	druid-hbase-emitter	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-skeleton	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-extensions	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-min-max	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-columns	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-emitter	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-contrib	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-view-maintenance	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-view-selection	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-data-storage	0.17.1.7.1.4.0-203
	org.apache.druid.extensions	druid-hbase-distcache-contrib	0.17.1.7.1.4.0-203
GCS Connector	com.google.cloud.bigtable	gcs-connector	2.1.2.7.1.4.0-203
	com.google.cloud.bigtable	gcs-connector	2.1.2.7.1.4.0-203
	com.google.cloud.bigtable	gcs-connector	2.1.2.7.1.4.0-203
	com.google.cloud.bigtable	gcs-connector	2.1.2.7.1.4.0-203
	com.google.cloud.bigtable	gcs-connector	2.1.2.7.1.4.0-203
Apache Hadoop	org.apache.hadoop	hadoop-aliyun	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-annotations	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-archive-logs	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-archives	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-assemblies	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-auth	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-aws	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-azure	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-azure-datalake	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-build-tools	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-client	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-client-api	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-client-integration-tests	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-client-minicluster	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-client-runtime	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-cloud-storage	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-common	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-datajoin	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-distcp	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-extras	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-fs2img	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-gridmix	3.1.1.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-hdds-client	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-common	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-config	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-container-service	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-docs	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-hadoop-dependency-client	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-hadoop-dependency-server	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-hadoop-dependency-test	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-interface-admin	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-interface-client	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-interface-server	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-server-framework	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-server-scm	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-test-utils	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdds-tools	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-hdfs	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-hdfs-client	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-hdfs-httpfs	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-hdfs-native-client	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-hdfs-nfs	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-hdfs-rbf	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-kafka	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-kms	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-app	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-common	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-core	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-hs	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-hs-plugins	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-jobclient	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-nativetask	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-shuffle	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-client-uploader	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-mapreduce-examples	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-maven-plugins	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-minicluster	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-minikdc	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-nfs	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-openstack	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-client	0.5.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-ozone-common	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-csi	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-datanode	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-dist	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-filessystem	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-filessystem-common	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-filessystem-hadoop2	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-filessystem-hadoop3	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-filessystem-shaded	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-insight	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-integration-test	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-interface-client	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-network-tests	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-ozone-manager	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-recon	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-reconcodegen	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-s3gateway	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-tools	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-ozone-upgrade	0.5.0.7.1.4.0-203
	org.apache.hadoop	hadoop-resourceestimator	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-rumen	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-sls	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-streaming	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-tools-dist	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-api	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-applications-distributedshell	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-applications-unmanaged-am-launcher	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-client	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-common	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-registry	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-applicationhistoryservice	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-common	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-nodemanager	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-resourcemanager	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-router	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-sharedcachemanager	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-tests	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-timeline-pluginstorage	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-timelineservice	3.1.1.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-client	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-common	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-server-2	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-tests	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-server-web-proxy	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-services-api	3.1.1.7.1.4.0-203
	org.apache.hadoop	hadoop-yarn-services-core	3.1.1.7.1.4.0-203
	org.apache.hadoop	mini-chaos-tests	0.5.0.7.1.4.0-203
Apache HBase	org.apache.hbase	hbase-annotations	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-checkstyle	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-client	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-client-project	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-common	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-endpoint	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-examples	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-external-blockcache	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-hadoop-compat	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-hadoop2-compat	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-hbtop	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-http	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-it	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-mapreduce	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-metrics	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-metrics-api	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-procedure	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-protocol	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-protocol-shaded	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-replication	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-resource-bundle	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-rest	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-rsgroup	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-server	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-shaded-client	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-shaded-client-byo-hadoop	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-shaded-client-project	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-shaded-mapreduce	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-shaded-testing-util	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-shaded-testing-util-tester	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-shell	2.2.3.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.hbase	hbase-testing-util	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-thrift	2.2.3.7.1.4.0-203
	org.apache.hbase	hbase-zookeeper	2.2.3.7.1.4.0-203
	org.apache.hbase.contrib	hbase-kerberos-model	1.0.0.7.1.4.0-203
	org.apache.hbase.contrib	hbase-kerberos-proxy	1.0.0.7.1.4.0-203
	org.apache.hbase.contrib	hbase-spark	1.0.0.7.1.4.0-203
	org.apache.hbase.contrib	hbase-spark-kt	1.0.0.7.1.4.0-203
	org.apache.hbase.contrib	hbase-spark-protocol	1.0.0.7.1.4.0-203
	org.apache.hbase.contrib	hbase-spark-protocol-shaded	1.0.0.7.1.4.0-203
	org.apache.hbase.fileformats	hbase-compat	1.0.0.7.1.4.0-203
Apache Hive	org.apache.hive	hive-accumulo-handler	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-beeline	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-classification	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-cli	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-common	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-contrib	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-druid-handler	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-exec	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-hbase-handler	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-hplsql	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-jdbc	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-jdbc-handler	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-kryo-registrator	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-kudu-handler	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-llap-client	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-llap-common	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-llap-ext-client	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-llap-server	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-llap-tez	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-metastore	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-pre-upgrade	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-serde	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-service	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-service-rpc	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-shims	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-spark-client	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-standalone-metastore	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-storage-api	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-streaming	3.1.3000.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.hive	hive-testutils	3.1.3000.7.1.4.0-203
	org.apache.hive	hive-vector-code-gen	3.1.3000.7.1.4.0-203
	org.apache.hive	kafka-handler	3.1.3000.7.1.4.0-203
	org.apache.hive.hcatalog	hive-hcatalog-core	3.1.3000.7.1.4.0-203
	org.apache.hive.hcatalog	hive-hcatalog-pig-adapter	3.1.3000.7.1.4.0-203
	org.apache.hive.hcatalog	hive-hcatalog-server-extensions	3.1.3000.7.1.4.0-203
	org.apache.hive.hcatalog	hive-hcatalog-streaming	3.1.3000.7.1.4.0-203
	org.apache.hive.hcatalog	hive-webhcat	3.1.3000.7.1.4.0-203
	org.apache.hive.hcatalog	hive-webhcat-java-client	3.1.3000.7.1.4.0-203
	org.apache.hive.shims	hive-shims-0.20	3.1.3000.7.1.4.0-203
	org.apache.hive.shims	hive-shims-0.23	3.1.3000.7.1.4.0-203
	org.apache.hive.shims	hive-shims-common	3.1.3000.7.1.4.0-203
	org.apache.hive.shims	hive-shims-scheduler	3.1.3000.7.1.4.0-203
Apache Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector_2.11	1.0.0.7.1.4.0-203
Apache Kafka	org.apache.kafka	connect	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-api	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-basic-auth-extension	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-file	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-json	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-mirror	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-mirror-client	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-runtime	2.4.1.7.1.4.0-203
	org.apache.kafka	connect-transforms	2.4.1.7.1.4.0-203
	org.apache.kafka	generator	2.4.1.7.1.4.0-203
	org.apache.kafka	jmh-benchmarks	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-clients	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-cloudera-plugins	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-examples	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-log4j-appender	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-examples	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-scala_2.11	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-scala_2.12	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-scala_2.13	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-test-utils	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-0100	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-0101	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-0102	2.4.1.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.kafka	kafka-streams-upgrade-system-tests-0110	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-10	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-11	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-20	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-21	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-22	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-streams-upgrade-system-tests-23	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka-tools	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka_2.11	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka_2.12	2.4.1.7.1.4.0-203
	org.apache.kafka	kafka_2.13	2.4.1.7.1.4.0-203
Apache Knox	org.apache.knox	gateway-adapter	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-admin-ui	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-applications	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-cloud-bindings	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-cm-integration	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-demo-ldap	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-demo-ldap-launcher	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-discovery-ambari	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-discovery-cm	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-docker	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-i18n	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-i18n-logging-log4j	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-i18n-logging-sl4j	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-ha	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-identity-assertion-common	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-identity-assertion-concat	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-identity-assertion-hadoop-groups	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-identity-assertion-pseudo	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-identity-assertion-regex	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-identity-assertion-switchcase	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-jersey	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-rewrite	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-rewrite-common	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-rewrite-func-hostmap-static	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-rewrite-func-inbound-query-param	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-rewrite-func-service-registry	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-rewrite-step-encrypt-uri	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-rewrite-step-secure-query	1.3.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.knox	gateway-provider-security-authc-anon	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-authz-acls	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-authz-composite	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-clientcert	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-hadoopauth	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-jwt	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-pac4j	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-preauth	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-shiro	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-provider-security-webappsec	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-release	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-server	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-server-launcher	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-server-xforwarded-filter	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-admin	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-as	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-definitions	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-hashicorp-vault	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-hbase	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-health	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-hive	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-idbroker	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-impala	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-jkg	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-knoxsso	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-knoxssout	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-knoxtoken	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-livy	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-metadata	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-nifi	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-nifi-registry	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-remoteconfig	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-rm	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-session	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-storm	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-test	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-tgs	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-vault	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-service-webhdfs	1.3.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.knox	gateway-shell	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-shell-launcher	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-shell-release	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-shell-samples	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-spi	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-test	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-test-idbroker	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-test-release-utils	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-test-utils	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-topology-simple	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-util-common	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-util-configinjector	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-util-launcher	1.3.0.7.1.4.0-203
	org.apache.knox	gateway-util-urltemplate	1.3.0.7.1.4.0-203
	org.apache.knox	hadoop-examples	1.3.0.7.1.4.0-203
	org.apache.knox	knox-cli-launcher	1.3.0.7.1.4.0-203
	org.apache.knox	knox-homepage-ui	1.3.0.7.1.4.0-203
	org.apache.knox	webhdfs-kerb-test	1.3.0.7.1.4.0-203
	org.apache.knox	webhdfs-test	1.3.0.7.1.4.0-203
Apache Kudu	org.apache.kudu	kudu-backup-tools	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-backup2_2.11	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-client	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-client-tools	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-hive	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-mapreduce	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-spark2-tools_2.11	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-spark2_2.11	1.12.0.7.1.4.0-203
	org.apache.kudu	kudu-test-utils	1.12.0.7.1.4.0-203
Apache Livy	org.apache.livy	livy-api	0.6.0.7.1.4.0-203
	org.apache.livy	livy-client-common	0.6.0.7.1.4.0-203
	org.apache.livy	livy-client-http	0.6.0.7.1.4.0-203
	org.apache.livy	livy-core_2.11	0.6.0.7.1.4.0-203
	org.apache.livy	livy-examples	0.6.0.7.1.4.0-203
	org.apache.livy	livy-integration-test	0.6.0.7.1.4.0-203
	org.apache.livy	livy-repl_2.11	0.6.0.7.1.4.0-203
	org.apache.livy	livy-rsc	0.6.0.7.1.4.0-203
	org.apache.livy	livy-scala-api_2.11	0.6.0.7.1.4.0-203
	org.apache.livy	livy-server	0.6.0.7.1.4.0-203
	org.apache.livy	livy-test-lib	0.6.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.livy	livy-thriftserver	0.6.0.7.1.4.0-203
	org.apache.livy	livy-thriftserver-session	0.6.0.7.1.4.0-203
Apache Lucene	org.apache.lucene	lucene-analyzers-common	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-icu	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-kuromoji	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-morfologik	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-nori	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-openssl	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-phonetic	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-smartcn	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-analyzers-stempel	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-backward-codecs	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-benchmark	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-classification	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-codecs	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-core	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-demo	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-expressions	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-facet	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-grouping	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-highlighter	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-join	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-memory	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-misc	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-monitor	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-queries	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-queryparser	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-replicator	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-sandbox	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-spatial	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-spatial-extras	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-spatial3d	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-suggest	8.4.1.7.1.4.0-203
	org.apache.lucene	lucene-test-framework	8.4.1.7.1.4.0-203
Apache Oozie	org.apache.oozie	oozie-client	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-core	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-distro	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-examples	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-fluent-job-api	5.1.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.oozie	oozie-fluent-job-client	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-server	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-distcp	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-git	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-hcatalog	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-hive	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-hive2	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-oozie	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-spark	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-sqoop	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-sharelib-streaming	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-tools	5.1.0.7.1.4.0-203
	org.apache.oozie	oozie-zookeeper-security-tests	5.1.0.7.1.4.0-203
	org.apache.oozie.test	oozie-mini	5.1.0.7.1.4.0-203
Apache ORC	org.apache.orc	orc-core	1.5.1.7.1.4.0-203
	org.apache.orc	orc-examples	1.5.1.7.1.4.0-203
	org.apache.orc	orc-mapreduce	1.5.1.7.1.4.0-203
	org.apache.orc	orc-shims	1.5.1.7.1.4.0-203
	org.apache.orc	orc-tools	1.5.1.7.1.4.0-203
Apache Phoenix	org.apache.phoenix	phoenix-client	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-core	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-hive	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-load-balancer	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-pherf	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-queryserver	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-queryserver-client	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-server	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-spark	5.0.0.7.1.4.0-203
	org.apache.phoenix	phoenix-tracing-webapp	5.0.0.7.1.4.0-203
Apache Ranger	org.apache.ranger	conditions-enrichers	2.0.0.7.1.4.0-203
	org.apache.ranger	credentialbuilder	2.0.0.7.1.4.0-203
	org.apache.ranger	embeddedwebserver	2.0.0.7.1.4.0-203
	org.apache.ranger	jisql	2.0.0.7.1.4.0-203
	org.apache.ranger	ldapconfigcheck	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-atlas-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-atlas-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-distro	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-examples-distro	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-hbase-plugin	2.0.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.ranger	ranger-hbase-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-hdfs-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-hdfs-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-hive-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-hive-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-intg	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kafka-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kafka-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kms	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kms-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kms-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-knox-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-knox-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kudu-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kylin-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-kylin-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-nifi-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-nifi-registry-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-ozone-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-ozone-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-plugin-classloader	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-plugins-audit	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-plugins-common	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-plugins-cred	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-plugins-installer	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-raz-adls	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-raz-hook-abfs	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-raz-intg	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-raz-processor	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-rms-common	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-rms-hive	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-rms-plugins-common	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-rms-webapp	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-sampleapp-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-schema-registry-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-solr-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-solr-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-sqoop-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-sqoop-plugin-shim	2.0.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.ranger	ranger-storm-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-storm-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-tagsync	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-tools	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-util	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-yarn-plugin	2.0.0.7.1.4.0-203
	org.apache.ranger	ranger-yarn-plugin-shim	2.0.0.7.1.4.0-203
	org.apache.ranger	sample-client	2.0.0.7.1.4.0-203
	org.apache.ranger	sampleapp	2.0.0.7.1.4.0-203
	org.apache.ranger	unixauthclient	2.0.0.7.1.4.0-203
	org.apache.ranger	unixauthservice	2.0.0.7.1.4.0-203
	org.apache.ranger	unixusersync	2.0.0.7.1.4.0-203
Apache Solr	org.apache.solr	solr-analysis-extras	8.4.1.7.1.4.0-203
	org.apache.solr	solr-analytics	8.4.1.7.1.4.0-203
	org.apache.solr	solr-cell	8.4.1.7.1.4.0-203
	org.apache.solr	solr-clustering	8.4.1.7.1.4.0-203
	org.apache.solr	solr-core	8.4.1.7.1.4.0-203
	org.apache.solr	solr-dataimporthandler	8.4.1.7.1.4.0-203
	org.apache.solr	solr-dataimporthandler-extras	8.4.1.7.1.4.0-203
	org.apache.solr	solr-jaegertracer-configurator	8.4.1.7.1.4.0-203
	org.apache.solr	solr-langid	8.4.1.7.1.4.0-203
	org.apache.solr	solr-ltr	8.4.1.7.1.4.0-203
	org.apache.solr	solr-prometheus-exporter	8.4.1.7.1.4.0-203
	org.apache.solr	solr-security-util	8.4.1.7.1.4.0-203
	org.apache.solr	solr-solrj	8.4.1.7.1.4.0-203
	org.apache.solr	solr-test-framework	8.4.1.7.1.4.0-203
	org.apache.solr	solr-velocity	8.4.1.7.1.4.0-203
Apache Spark	org.apache.spark	spark-avro_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-catalyst_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-core_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-graphx_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-hadoop-cloud_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-hive-thriftserver_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-hive_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-kubernetes_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-kvstore_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-launcher_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-mllib-local_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-mllib_2.11	2.4.0.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.spark	spark-network-common_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-network-shuffle_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-network-yarn_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-repl_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-sketch_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-sql-kafka-0-10_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-sql_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-streaming-kafka-0-10_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-streaming-kafka-0-8-assembly_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-streaming-kafka-0-8_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-streaming_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-tags_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-unsafe_2.11	2.4.0.7.1.4.0-203
	org.apache.spark	spark-yarn_2.11	2.4.0.7.1.4.0-203
Apache Sqoop	org.apache.sqoop	sqoop	1.4.7.7.1.4.0-203
	org.apache.sqoop	sqoop-test	1.4.7.7.1.4.0-203
Apache Tez	org.apache.tez	hadoop-shim	0.9.1.7.1.4.0-203
	org.apache.tez	hadoop-shim-2.8	0.9.1.7.1.4.0-203
	org.apache.tez	tez-api	0.9.1.7.1.4.0-203
	org.apache.tez	tez-aux-services	0.9.1.7.1.4.0-203
	org.apache.tez	tez-common	0.9.1.7.1.4.0-203
	org.apache.tez	tez-dag	0.9.1.7.1.4.0-203
	org.apache.tez	tez-examples	0.9.1.7.1.4.0-203
	org.apache.tez	tez-ext-service-tests	0.9.1.7.1.4.0-203
	org.apache.tez	tez-history-parser	0.9.1.7.1.4.0-203
	org.apache.tez	tez-javadoc-tools	0.9.1.7.1.4.0-203
	org.apache.tez	tez-job-analyzer	0.9.1.7.1.4.0-203
	org.apache.tez	tez-mapreduce	0.9.1.7.1.4.0-203
	org.apache.tez	tez-protobuf-history-plugin	0.9.1.7.1.4.0-203
	org.apache.tez	tez-runtime-internals	0.9.1.7.1.4.0-203
	org.apache.tez	tez-runtime-library	0.9.1.7.1.4.0-203
	org.apache.tez	tez-tests	0.9.1.7.1.4.0-203
	org.apache.tez	tez-yarn-timeline-cache-plugin	0.9.1.7.1.4.0-203
	org.apache.tez	tez-yarn-timeline-history	0.9.1.7.1.4.0-203
	org.apache.tez	tez-yarn-timeline-history-with-acls	0.9.1.7.1.4.0-203
	org.apache.tez	tez-yarn-timeline-history-with-fs	0.9.1.7.1.4.0-203
Apache Zeppelin	org.apache.zeppelin	sap	0.8.2.7.1.4.0-203
	org.apache.zeppelin	spark-interpreter	0.8.2.7.1.4.0-203

Project	groupId	artifactId	version
	org.apache.zepplin	spark-scala-2.11	0.8.2.7.1.4.0-203
	org.apache.zepplin	spark-shims	0.8.2.7.1.4.0-203
	org.apache.zepplin	spark2-shims	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-alluxio	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-angular	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-bigquery	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-cassandra_2.10	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-display	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-elasticsearch	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-file	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-flink_2.10	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-groovy	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-hbase	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-ignite_2.10	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-interpreter	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-jdbc	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-jupyter	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-kylin	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-lens	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-livy	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-markdown	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-neo4j	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-pig	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-python	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-scio_2.10	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-server	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-shell	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-spark-dependencies	0.8.2.7.1.4.0-203
	org.apache.zepplin	zeppelin-zengine	0.8.2.7.1.4.0-203
Apache ZooKeeper	org.apache.zookeeper	zookeeper	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-client-c	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-contrib-loggraph	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-contrib-rest	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-docs	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-jute	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-recipes-election	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-recipes-lock	3.5.5.7.1.4.0-203
	org.apache.zookeeper	zookeeper-recipes-queue	3.5.5.7.1.4.0-203

What's New In Cloudera Runtime 7.1.4

You must be aware of the additional functionalities and improvements to features of components in Cloudera Runtime 7.1.4. Learn how the new features and improvements benefit you.

What's New in Apache Atlas

Learn about the new features of Atlas in Cloudera Runtime 7.1.4.

Integrating Ozone with Apache Atlas

Integrating Ozone with Atlas, entities like Hive, Spark process, and NiFi flows, when created with Ozone path, results in Atlas creating Ozone path entities. For more information, see [Integrating Ozone with Atlas](#).

Restricting Atlas classifications based on user permission

The Atlas authorisation model has been updated with a new permission, type-read, to enable restricting visibility of types to specific users, groups, and roles. For more information, see [Restricting Atlas Classifications for users](#).

What's New in Cruise Control

There are no new features for Cruise Control in Cloudera Runtime 7.1.4

What's new in Data Analytics Studio

There are no new features for Data Analytics Studio in Cloudera Runtime 7.1.4.

What's New in Apache HBase

There are no new features for HBase in Cloudera Runtime 7.1.4

What's New in Apache Hadoop HDFS

There are no new features for Hadoop HDFS in Cloudera Runtime 7.1.4

For more information about HDFS, see [HDFS Overview](#)

What's New in Apache Hive

Learn about the new features for Hive in Cloudera Runtime 7.1.4.

After you upgrade to CDP and migrate old tables, you might want to briefly switch to Hive legacy behavior. Legacy behavior might solve compatibility problems with your scripts during data migration, for example, when running ETL. For more information, see [Configuring legacy CREATE TABLE behavior](#).

What's New in Hue

There are no new features for Hue in Cloudera Runtime 7.1.4.

What's New in Apache Impala

There are no new features for Impala in Cloudera Runtime 7.1.4

What's New in Apache Kafka

Learn about the new features of Kafka in Cloudera Runtime 7.1.4.

The kafka-clients library is no longer dependant on Apache Shiro and libpam4j

The LdapPlainServerCallbackHandler and PamPlainServerCallbackHandler custom Cloudera plugins introduced undesired dependencies for users of the kafka-client library. These plugins are now moved to a new module in Kafka that is only used on the broker side. As a result, the kafka-clients library is no longer dependent on Apache Shiro and libpam4j.

What's New in Apache Knox

There are no new features for Knox in Cloudera Runtime 7.1.4

What's New in Apache Kudu

Learn about the new features of Kudu in Cloudera Runtime 7.1.4.

RaftConsensus::DumpStatusHtml() does not block Raft consensus activity

kudu::consensus::RaftConsensus::CheckLeadershipAndBindTerm() needs to take the lock to check the term and the Raft role. When many RPCs come in for the same tablet, the contention can hog service threads and cause queue overflows on busy systems. With this improvement, RaftConsensus::DumpStatusHtml() no longer blocks Raft consensus activity and is not blocked by it either.

Adding --ignore_nonexistent for 'local_replica delete'

--ignore_nonexistent flag was added for the 'local_replica delete' tool. This makes the real-world scripting scenarios easier if trying to clean up tablet servers of particular tablet replicas.

Multiple tablet ids in 'local_replica delete'

The 'local_replica delete' tool allows multiple tablet identifiers to be specified and processed at once. This helps to reduce the overall latency caused by that opening tablet server's metadata takes significant time.

Avoid an epoll cycle when data can be sent immediately

If the transfer queue was empty prior to a transfer is enqueued, immediate writing is attempted. If that attempt fails to fully write, epoll watching is enabled.

Avoid calling Schema::find_column() once per RowBlock in columnar serialization

When a row block is being serialized in the columnar format the projection calculation happens once per Scan RPC and per-rowblock calls.

What's New in Apache Oozie

There are no new features for Oozie in Cloudera Runtime 7.1.4

What's New in Apache Hadoop Ozone

There are no new features for Apache Hadoop Ozone in Cloudera Runtime 7.1.4.

What's New in Apache Phoenix

There are no new features for Phoenix in Cloudera Runtime 7.1.4

What's New in Schema Registry

Learn about the new features for Schema Registry in Cloudera Runtime 7.1.4

Security-related header controls in Schema Registry responses

The following headers appear in Schema Registry HTTP responses to provide security-related restrictions and instructions to the browser:

- Cache-Control
- Content-Security-Policy
- Content-Type-Options
- Strict-Transport-Security (if SSL is enabled)
- X-Frame options
- XSS-Protection

What's New in Cloudera Search

There are no new features for Cloudera Search in Cloudera Runtime 7.1.4

What's New in Sqoop

There are no new features for Sqoop in Cloudera Runtime 7.1.4.

To access the latest Sqoop documentation on Cloudera's documentation 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 Replication Manager

There are no new features for Streams Replication Manager in Cloudera Runtime 7.1.4

What's new in Streams Messaging Manager

Learn about the new features of Streams Messaging Manager in Cloudera Runtime 7.1.4.

Security-related header controls in SMM REST API server responses

The following HTTP headers appear in SMM REST API HTTP responses to provide security-related restrictions and instructions to the browser:

- Strict-Transport-Security
- Cache-Control

Security-related header controls in SMM UI server responses

The following HTTP headers appear in SMM UI server HTTP responses to provide security-related restrictions and instructions to the browser:

- Cache Control
- XSS-PROTECTION
- Content-Security-Policy
- Strict-Transport-Security
- X-Content-Type-Options
- X-Frame-Options

Disabled TLS 1.0 and TLS 1.1 in SMM UI

TLS 1.0 and TLS 1.1 are now disabled in SMM UI.

What's New in Apache Hadoop YARN

Learn about the new features of Hadoop YARN in Cloudera Runtime 7.1.4.

Resource Allocation - Absolute Mode

You can now allocate resources to queues using Absolute Mode. You can specify the number of units of vCores and memory resources by using the Absolute allocation mode.

For more information about resource allocation mode, see [Resource Allocation Overview](#).

What's New in Apache ZooKeeper

There are no new features for ZooKeeper in Cloudera Runtime 7.1.4.

Fixed Issues In Cloudera Runtime 7.1.4

You can review the list of reported issues and their fixes in Cloudera Runtime 7.1.4.

Fixed Issues in Apache Atlas

Review the list of Atlas issues that are resolved in Cloudera Runtime 7.1.4.

ATLAS-3806: Atlas notifications to Ranger are missing propagated classifications

When an entity is updated or created, Atlas correctly propagated classifications from the parent table or tables to the new entity. However, when Atlas notified Ranger of the new table, the notification did not include the propagated classification. If Ranger configuration included a tag-based access policy that corresponds to the Atlas classification, the policy was not applied to the new table. For example, if you marked a table with a classification to indicate that it had sensitive data (such as "PII"), then used fields from that table to create another table in a CTAS operation, Atlas propagated the PII classification from the parent table to the new table. The data Atlas subsequently sent to Ranger did not have the propagated "PII" classification, and therefore Ranger did not apply the tag-based access policy to the table. This problem is fixed in this release.

CDPD-14877: SAC creates ozone entities as hdfs_path entities.

This issue is now resolved.

CDPD-14031: In the Spark Atlas Connector, few S3 entities are created using the V1 S3 model instead of the updated V2 S3 model.

Use Atlas S3 v2 models in Spark Atlas Connector. This issue is now resolved.

CDPD-11790: Shell entity is not resolved to the Complete entity under certain conditions.

Shell entities with duplicate qualifiedName are no longer created when processing message from Spark Atlas Connector. The issue is addressed by this implementation and

[ATLAS-3874](#)

Fixed issues in Data Analytics Studio

There are no fixed issues for Data Analytics Studio in Cloudera Runtime 7.1.4.

Fixed Issues in Apache Hadoop

Review the list of Hadoop issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-7383: The ABFS and ADL connectors compatible with Alpine Linux and other platforms which have libssl1.1-1.1.1b-r1 as their native OpenSSL implementation

See [HADOOP-16460](#) and [HADOOP-16438](#)

Fixed Issues in HDFS

Review the list of HDFS issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-17842: MAPREDUCE-7298: For SSL Enabled Service, the distcp job is not closed leading to leaked TrustStore Reloader Thread.

This issue is now resolved.

CDPD-12296: HDFS-15313: If any Inodes are referenced in the active file system tree, the Inodes are deleted when a snapshot is deleted.

This issue is now resolved. Now, Inodes are not deleted when a snapshot is deleted.

CDPD-6100

This improvement makes HDFS NameNode leave safemode immediately if all blocks have reported in, reducing the cluster startup time.

CDPD-2946: Slow reading and writing of erasure-coded files

The ISA-L library is not packaged with HDFS as a result of which HDFS erasure coding falls back to the Java implementation which is much slower than the native Hadoop implementation. This slows down the reading and writing of erasure-coded files.

Fixed Issues in Apache HBase

Review the list of HBase issues that are resolved in Cloudera Runtime 7.1.4.

OPSAPS-57422: HBase 2.0 JMX GET metrics changed.

Metrics removed in HBase2 are no longer polled by Cloudera Manager. This issue is now resolved.

Fixed Issues in Apache Hive

Review the list of Hive issues that are resolved in Cloudera Runtime 7.1.4.

Fixed issues running Hive ACID when using Cloudera Runtime with Hive Virtual Warehouses

Hive ACID is supported for Hive Virtual Warehouses using Cloudera Runtime 7.1.4 in CDP Private Cloud Base.

CDPD-13345: In jdbcUrl add configuration to create External + purge table by default.

This issue is now resolved.

CDPD-16123: Fixed the problem preventing spin off of a separate Hive on Tez session for the Hive Warehouse Connector.

CDPD-16040: Fixed the problem requiring the --hs2-url option in sqoop import.

The default JDBC URL is set to the HS2 URL automatically.

CDPD-12301: Fixed Spark Hive Streaming issue that caused HWC to fail and throw a java.lang.NoSuchMethodException message.

CDPD-9214: Fixed absence of query progress indicators when Spark-Hive queries through the Hive JDBC connector include a where clause.

CDPD-9214: Fixed the problem identifying the Tez YARN application ID or IDs run by a Spark session.

CDPD-9297: Fixed the issue with DESCRIBE statements from SparkSQL to fail.

CDPD-6761: Fixed how Apache Hive ignores the database location provided by Apache Spark.

OPSAPS-57198: Fixed the issue causing Erasure coding warnings to appear Cloudera Manager even though you were not using erasure coding.

Compaction issues fixed in this release

CDH-84872: HIVE-22705: LLAP cache is polluted by query-based compactor

CDPD-12497: HIVE-22971: Eliminate file rename in insert-only compactor

CDPD-12576: HIVE-23531: Major CRUD QB compaction failing with ClassCastException when vectorization off

CDPD-12577: HIVE-23703: Major QB compaction with multiple FileSinkOperators results in data loss and one

CDPD-13000: HIVE-23555: Cancel compaction jobs when hive.compactor.worker.timeout is reached

CDPD-13000: HIVE-23591: When Worker fails to connect to metastore it should wait before retrying

CDPD-14959: HIVE-23825: Create a flag to turn off _orc_acid_version file creation

CDPD-15116: HIVE-23853: CRUD based compaction also should update ACID file version metadata

CDPD-15884: HIVE-23966: Minor query-based compaction always results in delta directories with minWriteId=1

CDPD-15955: HIVE-23929: Run compaction as partition owner

CDPD-15956: HIVE-22255: Hive doesn't trigger Major Compaction automatically if table contains only base files

CDPD-16031: HIVE-24001: Don't cache MapWork in tez/ObjectCache during query-based compaction

CDPD-16132: HIVE-24024: Improve logging around CompactionTxnHandler

CDPD-16149: HIVE-24015: Disable query-based compaction on MR execution engine

CDPD-16183: HIVE-24023: Hive parquet reader can't read files with length=0

CDPD-16298: HIVE-23763: Query based minor compaction produces wrong files when rows with different buckets Ids are processed by the same FileSinkOperator

CDPD-16471: HIVE-24089: Run QB compaction as table directory user with impersonation

CDPD-16578: HIVE-24096: Abort failed compaction's txn on TException or IOException

Hive snapshot isolation issues fixed in this release

CDPD-17113: HIVE-23032: Add batching in Lock generation

CDPD-17114: HIVE-22750: Consolidate LockType naming

CDPD-17108: HIVE-23349: ACID: Concurrent MERGE INSERT operations produce duplicates

CDPD-17112: HIVE-24000: Put exclusive MERGE INSERT under the feature flag

CDPD-17109: HIVE-23503: ValidTxnManager doesn't consider txns opened and committed between snapshot generation and locking when evaluating ValidTxnListState

CDPD-11683: HIVE-23311: Fix ValidTxnManager regression

CDPD-16308: HIVE-24125: Incorrect transaction snapshot invalidation with unnecessary writeset check for exclusive operations

Fixed Issues in Hue

There are no fixed issues for Hue in Cloudera Runtime 7.1.4.

Fixed Issues in Apache Impala

Review the list of Impala issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-10444: Update the version of Atlas used by Impala

Resolves the inconsistency between the versions of jackson-databind used by Atlas and Impala so that you no longer see the exception `NoClassDefFoundError` after the call to `QueryEventHookManager#executeQueryCompleteHooks()` in Impala.

CDPD-14548 and CDPD-13701: IMPALA-10060: Update PostgreSQL JDBC Driver (pgjdbc) for Impala (CVE-2020-13692).

Upgrades the library, PostgreSQL JDBC Driver (pgjdbc). The previous version 42.2.5 was affected with vulnerability CVE-2020-13692. It is upgraded to version 42.2.14. This issue is now resolved.

IMPALA-77: Impala does not properly handle HDFS viewfs mounts (federated cluster) - Queries fail with "wrong filesystem" error

This issue is fixed in 1.0 by improving the error message.

Fixed Issues in Apache Kafka

Review the list of Kafka issues that are resolved in Cloudera Runtime 7.1.4.

KAFKA-9839: IllegalStateException on metadata update when broker learns about its new epoch after the controller

This is a backported fix. The broker now accepts control requests if the broker epoch is newer than the current epoch.

OPSAPS-57907: The Kafka metric collector adapter generates high CPU load

This issue is now resolved.

OPSAPS-57113: The Kafka Broker Advanced Configuration Snippet (Safety Valve) for ssl.properties does not propagate configurations correctly

Configuration containing dollar signs is now propagated correctly.

CDPD-17172: Fix of consumer group rebalance issue KAFKA-9752 backported.

This issue is now resolved.

Kafka Connect fails to communicate with secured Schema Registry

The `*.converter.sasl.jaas.config` connector property can now be used to specify the JAAS configuration required when connecting to a secured Schema Registry server. For example:

```
"value.converter.sasl.jaas.config": "com.sun.security.auth.module.Krb5LoginModule
    required useKeyTab=true storeKey=true keyTab=\"/path
ToKeytab/fileName.keytab\"
    principal=\"principal@REALM.COM\" ; ",
```

If multiple Kafka Connect roles are installed, you must provide a JAAS configuration that uses a headless principal and a keytab. If a valid configuration is added, Kafka Connect will successfully communicate with the server. Creating a JAAS configuration with a RegistryClient entry and adding the file to the Kafka environment is no longer required.

Fixed Issues in Apache Kudu

Review the list of Kudu issues that are resolved in Cloudera Runtime 7.1.4.

KUDU-3152: KuduPredicate class in Java client does not handle Date columns

Prior to this fix, if you had a table with DATE column, you could not scan for it using the java client. A check for minimum and maximum boundaries of integer representation of `java.sql.Date` was added to match `MIN_DATE_VALUE` and `MAX_DATE_VALUE` in `DateUtil`.

Fixed Issues in Apache Oozie

Review the list of Oozie issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-57814: Cloudera Manager will populate the newly introduced ZooKeeper SSL fields for Oozie when AutoTLS is enabled on the Cluster and SSL is enabled for ZooKeeper.

CDPD-11967: Fix to address CWE-693: Protection Mechanism Failure

CDPD-12742: Oozie is unable to communicate with ID Broker and fails to obtain a delegation token because of a missing Jar.

This issue is now resolved.

CDPD-13702: Fix to address CVE-2020-13692

CDPD-14600: Fix to address CVE-2016-3088

CDPD-14964: When a Java action calls `System.exit()`, an exception is logged by Oozie for Sqoop actions even after successful workflow actions.

This issue is now resolved.

CDPD-10746: Fix to address CVE-2019-17571

CDPD-9895: Various errors when trying to use an S3 filesystem

Oozie is now fully compatible with S3.

CDPD-9761: There is a sub workflow run in independent mode that runs a fork action which contains two (or more) actions

These actions inside the fork action run in parallel mode, and they have some seconds delay in between them. If a parameter is passed to one of these actions, that cannot be resolved, then it changes its status to FAILED, and also the workflows state to FAILED. The other actions state which are not started yet will stuck in PREP state forever. The correct behaviour would be to KILL the remaining actions as well as the workflow. Note: this bug only occurs when it is run in independent mode. If it has a parent workflow, then the parent workflow will kill this workflow after 10 minutes because of the callback process.

CDPD-9721: Upgrade built-in spark-hive in Oozie

Oozie is using the Spark-Hive library from the stack.

CDPD-9220: Oozie spark actions using `--keytab` fail due to duplicate dist. cache

Oozie spark actions add everything in the distributed cache of the launcher job to the distributed cache of the spark job, meaning the keytab is already there, then the `--keytab` argument tries to add it again causing the failure.

CDPD-9189: Apache Pig support was completely removed from Oozie

CDPD-7108: In case we have a workflow which has, lets say, 80 actions after each other, then the validator code "never" finishes

CDPD-7107: The following were added to the spark opts section of the spark action: `--conf spark`

CDPD-7106: query tag is not functional for Hive2 action node in oozie

Workflow is intended to create a hive table using Hive2 action node. Though workflow run successfully, table is not created.

CDPD-7105: Oozie workflow processing becomes slow after the increase of rows in WF_JOBS and WF_ACTIONS tables when running against SQL Server

CDPD-6877: When you create a MapReduce action which then creates more than 120 counters, an exception was thrown

CDPD-6630: Oozie by default gathers delegation tokens for the nodes defined in MapReduce

CDPD-5168: Logging enhancements in CoordElFunctions for better supportability

CDPD-4826: Oozies web server does not work when TLS is enabled and Open JDK 11 is in use

This issue is now fixed.

Fixed issues in Ozone

Review the list of Ozone issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-17352: HDDS-4193: Range used by AWS S3 multipartUpload copy-from-source must be inclusive.

This issue is now resolved.

CDPD-15072:

Ozone filesystem trash is not automatically cleaned up.

This issue is now resolved. This fix disables the Hadoop filesystem trash support in Ozone. Usually, when a file stored on a Hadoop compatible filesystem is deleted via the fs shell, the file is moved to the trash and eventually purged after the fs.trash.interval has elapsed. When the target filesystem is Ozone, the file is deleted immediately without being moved to trash. Trash support for the Ozone filesystem will be added in the subsequent release of CDP Private Cloud Base.

CDPD-15243: The Ozone volume s3v is an internal volume that is used to contain all buckets created through the s3 gateway.

When you create the first s3 bucket, the s3 volume is created automatically with the appropriate ACLs on a secure cluster.

CDPD-15080: Ozone key names are now interpreted as filesystem paths and are normalized before storing in Ozone.

Using the Hadoop filesystem API, you can now access the keys uploaded via the Ozone S3 API.

CDPD-7358: Upgrade to Guava 28.1 to avoid CVE-2018-10237

Ozone has been upgraded to use Guava version 28.1 to avoid CVE-2018-10237.

CDPD-7370: Ozone - Upgrade to Jetty 9.4.26 to avoid CVEs

Ozone now uses Jetty 9.4.26, which addresses the following CVEs: CVE#2017#7656, CVE#2017#7657, CVE#2017#7658, CVE#2018#12536, CVE#2017#9735, CVE#2019#10247.

Fixed Issues in Phoenix

There are no fixed issues for Phoenix in Cloudera Runtime 7.1.4.

Fixed Issues in Apache Ranger

Review the list of Ranger issues that are resolved in Cloudera Runtime 7.1.4.

ENGESC-6380: CDPD-15729 fix is backported.

This issue is resolved.

CDPD-17756: Normalize naming and processing of configuration parameters for chained plugin.

Chained plugin configuration must copy selected configuration parameters from root plugin's configuration and rename them to set up for creation and operation of chained plugin. This issue is now resolved.

CDPD-15715: The Zone Name field is not populated in ranger audits when there is a deny policy in the security zone or if there is no policy granting access to the matching resource path in the security zone.

This issue is now resolved.

CDPD-15401: When you enable Hive Metastore lookup in Ranger admin, resource lookup returns nothing and displays an error.

This issue is now resolved. You must use this step as a solution: `sudo ln -s /opt/cloudera/parcels/*<CDH-version>*/jars/libfb303-0.9.3.jar /opt/cloudera/parcels/*<CDH-version>*/lib/ranger-admin/ews/webapp/WEB-INF/lib/libfb303-0.9.3.jar`

CDPD-14269 and CDPD-14289: Failed resource filtering in Ranger Policy Export.

Exporting tag policies result in a 204 error when the polResource query parameter is used.

CDPD-12848: When you try to create multiple policies using the API having same non-existing group, the group creation fails due to multiple threads trying to create the same group at once.

Separate threads are now created for retry group creation and checks if the group is previously created and associate it with policy.

CDPD-10072: Ranger Ozone plugin unable to write to solr audits in SSL enabled cluster

This issue is now resolved. A separate folder libext is added under the Ozone library path and all the ranger plugin jars are added under this new folder.

Fixed Issues in Schema Registry

Review the list of Schema Registry issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-15097: Schemas with null description do not show up in the UI

Previously, when using the UI or API calls, the search results would sometimes list schemas according to description instead of name matches. Schemas without a description would not appear in the results.

In this release, schemas with a null description are now listed correctly and the UI search results appear correctly.

OPSAPS-58007 and OPSAPS-55940: Configuring the JDBC connector jar file location for Schema Registry and Streams Messaging Manager

Previously, in order for the start-up script to find the Schema Registry and SMM JDBC-connector jar files, you had to download the jar files and place them in the correct location.

Now the connector jar files are in the default path according to the configured database type.

OPSAPS-57467: Hard-coded parcel directory causes failure when non-standard path is used and TLS is enabled

Previously, the parcel directory path was hard-coded in the Schema Registry scripts and caused failures when a non-standard path was used and TLS was enabled.

The parcel directory path is no longer hard-coded in the Schema Registry CSD.

Fixed Issues in Cloudera Search

Review the list of Search issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-16332

Solr - Backport fixes for CVE-2020-13941 - SOLR-14515/SOLR-14561

Fixed Issues in Spark

Review the list of Spark issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-2650: Spark can't write ZSTD and LZ4 compressed Parquet to dynamically partitioned table.

This issue is resolved.

CDPD-3783: Unable to create database in spark.

This issue is resolved.

CDPD-12541: If an insert statement specifies partition statically and dynamically, there is a chance for data loss. To prevent data loss, Spark service displays an exception error.

This issue is now resolved. Spark allows you to specify a mix of static and dynamic partitions in an insert statement.

CDPD-14203: Introduced the `spark.executor.rpc.bindToAll` property to support multihoming when the new configuration is set to true bind to 0.0.0.0. The configuration defaults to false.

This issue is now resolved.

CDPD-10532: Update log4j to address CVE-2019-17571

Replaced log4j with an internal version to fix CVE-2019-17571.

CDPD-10515: Incorrect version of jackson-mapper-asl

Use an internal version of jackson-mapper-asl to address CVE-2017-7525.

CDPD-7882: If an insert statement specifies partitions both statically and dynamically, there is a potential for data loss

To prevent data loss, this fix throws an exception if partitions are specified both statically and dynamically. You can follow the workarounds provided in the error message.

CDPD-15773: Spark/Hive interaction causes deadlock.

In the previous versions, applications that share a Spark Session across multiple threads experienced a deadlock accessing the HMS. This issue is now resolved.

CDPD-14906: Spark reads or writes TIMESTAMP data for values before the start of the Gregorian calendar. This happens when Spark is:

- Using dynamic partition inserts.
- Reading or writing from an ORC table when `spark.sql.hive.convertMetastoreOrc=false` (the default is true).
- Reading or writing from an Orc table when `spark.sql.hive.convertMetastoreOrc=true` but `spark.sql.orc.impl=hive` (the default is native).
- Reading or writing from a Parquet table when `spark.sql.hive.convertMetastoreParquet=false` (the default is true).

This issue is now resolved.

Fixed Issues in Apache Sqoop

Review the list of Sqoop issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-12646: Sqoop does not close the open database connection before submitting the MapReduce Job. The open connection utilizes resources and displays an error message in the log when the connection times out.

This issue is resolved.

CDPD-16245: Sqoop uses the Hive column mapping even when you imported to HDFS.

Now, it uses the Hive column mapping for Hive imports and the Java column mapping for others.

This issue is resolved.

CDPD-16040: SGOOP-3458: When Sqoop is invoked without the `--hs2-url` parameter, it fails because the beeline command line script is unable to find the Hadoop installation.

This issue is resolved.

Fixed Issues in Streams Replication Manager

Review the list of Streams Replication Manager issues that are resolved in Cloudera Runtime 7.1.4.

CSP-956: Topics or groups added to white or blacklists are not returned when using `srn-control --list`

The `srn-control` tool and the SRM driver are now able to read the full white and blacklist from the `srn-control.<alias>.internal` configuration topic.

Fixed Issues in Streams Messaging Manager

Review the list of Streams Messaging Manager issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-12981: Security related header control is not available in Streams Messaging Manager.

Added Security related HTTP headers to Streams Messaging Manager UI server responses: Cache Control, XSS-PROTECTION, Content-Security-Policy, Strict-Transport-Security, X-Content-Type-Options, and X-Frame-Options.

CDPD-15552: Streams Messaging Manager throws `ConsumerGroups` exception.

Streams Messaging Manager tries to parse Kafka connectors as ConsumerGroups and throws the following exception:

```
Error reading field 'X': Error reading string of length....
```

OPSAPS-57410: Add Security related headers to the Streams Messaging Manager Rest API Server responses.

Strict-Transport-Security and Cache-Control are added. This issue is now resolved.

CDPD-14930: Streams Messaging Manager displays incorrect values of metrics.

For certain time periods, metrics such as Bytes In, Bytes out, Messages In are displaying incorrect value. This is due to CM metrics rollup where granularity of more than 1 min causes metrics to lose metric information resulting in incorrect calculation.

CDPD-14544: Streams Messaging Manager displays smaller topic metric values for larger time.

On the overview page, smaller topic metric values were displayed for larger time periods in comparison to the smaller time periods due to larger rollup intervals.

CDPD-14341: Visibility of metrics fetching failure needs improvement.

When Streams Messaging Manager fails to fetch metrics from CM, the failure is not visible on the UI and it is difficult to determine if there are any fetching errors in the logs.

Now, if Streams Messaging Manager fails to fetch metrics from CM, this failure will be saved and displayed when the user accesses the Streams Messaging Manager UI.

CDPD-10771: CPU usage charts showing incorrect values.

The CPU Usage charts for brokers on Streams Messaging Manager UI is showing incorrect values in CM.

OPSAPS-58007: Manually copy database driver libraries.

Users need to manually copy database driver libraries to the lib folders of Streams Messaging Manager.

Now, Streams Messaging Manager can use the database driver jars that are normally deployed to a shared location (typically: /usr/share/java).

OPSAPS-57745: CM displays incorrect role.

Cloudera Manager displays incorrect role status when the Streams Messaging Manager UI process fails or stops.

OPSAPS-57539: Streams Messaging Manager not usable after upgrade.

When upgrading the Cloudera Runtime, the previous Runtime's Streams Messaging Manager UI process was not being killed and making Streams Messaging Manager unusable.

Now, the Streams Messaging Manager UI stop script has been improved, so that it will kill the child processes to prevent the Streams Messaging Manager UI process from being stuck.

OPSAPS-57468: Hardcoded parcel directory causes failure.

Hardcoded parcel directory causes failure when non-standard path is used and TLS is enabled.

OPSAPS-55940: Connector jar location for Streams Messaging Manager changed.

Streams Messaging Manager used to come pre-packaged with the mysql jar file. This file needed to be located under the ./bootstrap/libs and ./libs directories in order for the startup script to find them. But the mysql connector JARs are not found in the libs folders anymore and you need to download the JARs and place them in the correct location manually.

Now, the fix adds configuration to use connector jars provided in default path according to the configured DB type.

Fixed Issues in Apache YARN

Review the list of YARN issues that are resolved in Cloudera Runtime 7.1.4.

OPSAPS-50291: "HADOOP_HOME,PATH,LANG,TZ" are now added by default to the yarn.nodemanager.env-whitelist Yarn configuration option.

This issue is resolved.

CDPD-16217: Clicking on queue in Capacity Scheduler, the legacy UI does not show any application.

This issue is now resolved.

YARN-10316:

The convert maxAppsDefault and maxRunningApps properties are not available for converting from Fair Scheduler to Capacity Scheduler.

Fixed Issues in Zeppelin

Review the list of Zeppelin issues that are resolved in Cloudera Runtime 7.1.4.

CDPD-10187: Incorrect version of jackson-mapper-asl.

Use internal version of jackson-mapper-asl to handle CVE-2017-7525.

CDPD-1683: Zeppelin demo users have been removed

Use cluster users to access Zeppelin. For information on provisioning users in CDP, see [Onboarding users](#).

CDPD-880, CDPD-1685: Shell, JDBC, and Spark interpreters have been removed

Use an available interpreter. For Spark functionality, use the Livy interpreter.

CDPD-3047: Markdown interpreter does not handle certain numbered list syntax correctly

Using the plus sign (+) or asterisk (*) to continue a numbered list using the %md interpreter results in bullet point entries instead.

Hotfixes in Cloudera Runtime 7.1.4

You can review the list of CDP Private Cloud Base hotfixes rolled into Cloudera Runtime 7.1.4. This will help you to verify if a hotfix provided to you on a previous CDP Private Cloud Base release was included in this release.

- 7.1.1-1.cdh7.1.1.p2002.5200993
- 7.1.1-1.cdh7.1.1.p2004.5453188

Known Issues In Cloudera Runtime 7.1.4

You must be aware of the known issues and limitations, the areas of impact, and workaround in Cloudera Runtime 7.1.4.



Note: CDSW does not support RPM-based installation on CDP Private Base. (RPM installation is deprecated and only supported on HDP and CDH 5. For CDH6 and onward, Cloudera recommends you to use CSD-based installations.)

Known Issues in Apache Atlas

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

CDPD-5542: AWS S3 Bulk and Incremental Extraction is currently not supported on 7.1.4.

None.

CDPD-17355: Atlas AWS extraction issue due to KeyError: 'entities'.

AWS S3 extraction does not happen as the extractor.sh is missing from the host.

None.

CDPD-14877:

In the Ozone Atlas integration, only the Spark-Atlas connector is failing.

Running the Spark query through the Ozone path, an incomplete entity is created.

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@58112bc4 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 blob locations (ABFS) doesn't have the clusterName appended to it as do hdfs_path entities in other location types.

None.

CDPD-5933, 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 "@" 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 Cruise Control

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

Cruise Control cannot collect metrics if Kafka topics have dot in their name

The Cloudera Manager metrics fetcher returns topic and topic-partition metrics with dot in their topic name field. Even though Kafka, by default, replaces the dots with underscore, the original topic name is restored in the process of Cloudera Manager metric collection. Because of this, Cruise Control cannot match the topic names with the default topic forms in Kafka, and causes failure in the metrics collection.

None.

Cruise Control might fail after CDP Private Cloud Base upgrade

When upgrading to a new version of CDP Private Cloud Base, during the restart Cruise Control might fail. Cruise Control needs metrics from Cloudera Manager to start up properly. After the upgrade, the metric collector might need more attempts to restart than it is configured in default. This can cause Cruise Control to fail after the upgrade.

1. In Cloudera Manager, go to Cruise Control > Configuration .
2. Search for Process Start Retry Attempts.
3. Set the value to 20 for Process Start Retry Attempts.

After setting the value, continue the upgrade after restarting Cruise Control.

Cruise Control might fail at first run

When you install Cruise Control either individually or using the Compute Cluster - StreamingMessaging(Full) deployment, Cruise Control might fail at the first run. This is caused by the difference between the Security Protocol in Kafka and in Cruise Control.

To avoid and solve this issue, see the [Add Cruise Control documentation](#).

Cruise Control capacity bootstrapping ignores deleted log directories

Log directories remain in the metrics database after a log directory is removed from Kafka. This causes Cruise Control unable to start up as it tries to query the metrics in Cloudera Manager without any data in them.

You need to stop the service monitor and delete the database (by default it can be found at: /var/lib/cloudera-service-monitor). Restart the service monitor and also Cruise Control.

Cruise Control does not package cruise-control-version.properties

The python client cannot be used as Cruise Control does not give any version information in HTTP response headers. In this version, Cruise Control does not support generating the cruise-control-version.properties file that is required by the python client for compatibility checks.

None

Some Kafka metrics are collected incorrectly in Cruise Control

Using the network inbound/outbound capacity goals can cause the network inbound metrics to increase in the logs.

Until the issue is resolved, avoid using the network inbound/outbound capacity goal.

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 execute 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.

Technical Service Bulletins

TSB 2022-581: Issues with “DAG ID” and “APP ID” visibility when exploring jobs in Data Analytics Studio

When using Data Analytics Studio (DAS) with Cloudera Data Platform (CDP) Private Cloud Base, sometimes the DAG ID and APP ID will not be visible to DAS.

Knowledge article:

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-581: Issues with “DAG ID” and “APP ID” visibility when exploring jobs in Data Analytics Studio](#)

Known Issues in Apache Hadoop

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

CDPD-10352: Hive on Tez cannot run certain queries on tables stored in encryption zones. This occurs when KMS connection is SSL encrypted and a self-signed certificate is used. You may see SSLHandshakeException in Hive logs in this case.

There are two workarounds: 1. You can install a self-signed SSL certificate into cacerts file in all hosts. 2. You can copy ssl-client.xml to a directory that is available in all hosts. Then you must set the `tez.aux.uris=path-to-ssl-client.xml` property in Hive on Tez advanced configuration.

Technical Service Bulletins

TSB 2021-434: KMS Load Balancing Provider Fails to invalidate Cache on Key Delete

The KMS Load balancing Provider has not been correctly invalidating the cache on key delete operations. The failure to invalidate the cache on key delete operations can result in the possibility that data can be leaked from the framework for a short period of time based on the value of the `hadoop.kms.current.key.cache.timeout.ms` property. Its default value is 30,000ms. When the KMS is deployed in an HA pattern the `KMSLoadBalancingProvider` class will only send the delete operation to one KMS role instance in a round-robin fashion. The code lacks a call to invalidate the cache across all instances and can leave key information including the metadata and key stored (the deleted key) in the cache on one or more KMS instances up to the key cache timeout.

Upstream JIRA

- [HADOOP-17208](#)
- [HADOOP-17304](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2020-434: KMS Load Balancing Provider Fails to invalidate Cache on Key Delete](#)

Known Issues in Apache HBase

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

HBASE-24885: If an operator uses HBCK2 to invoke multiple `assigns` operations against one Region or happens to invoke HBCK2 `assigns` while HBase is re-assigning a Region, it is possible that the Region will be abnormally assigned. For example, unassigned, stuck in transition, and doubly-assigned.

Obtain a fix for this issue. Operators should definitely not schedule multiple assigns for a single Region at the same time, however there is still a potential race condition.

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

Stop HBase using Cloudera Manager before deleting an OpDB Data Hub cluster.

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

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

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

HDFS encryption with HBase

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

None

AccessController postOperation problems in asynchronous operations

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

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

None

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

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

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

Storing Medium Objects (MOBs) in HBase is currently not supported

Storing MOBs in HBase relies on bulk loading files, and this is not currently supported when HBase is configured to use cloud storage (S3).

Workaround: None

Technical Service Bulletins

TSB 2021-453: Snapshot and cloned table corruption when original table is deleted

HBASE-25206 can cause data loss either through corrupting an existing hbase snapshot or destroying data that backs a clone of a previous snapshot.

Upstream JIRA

[HBASE-25206](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-453: HBASE-25206 "snapshot and cloned table corruption when original table is deleted"](#)

TSB 2021-463: Snapshot and cloned table corruption when original table is deleted

The HDFS short-circuit setting `dfs.client.read.shortcircuit` is overwritten to disabled by `hbase-default.xml`. HDFS short-circuit reads bypass access to data in HDFS by using a domain socket (file) instead of a network socket. This alleviates the overhead of TCP to read data from HDFS which can have a meaningful improvement on HBase performance (as high as 30-40%).

Users can restore short-circuit reads by explicitly setting `dfs.client.read.shortcircuit` in HBase configuration via the configuration management tool for their product (e.g. Cloudera Manager or Ambari).

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-463: HBase Performance Issue](#)

TSB 2021-494: Accumulated WAL Files Cannot be Cleaned up When Using Phoenix Secondary Global Indexes

The Write-ahead-log (WAL) files for Phoenix tables that have secondary global indexes defined on them, cannot be automatically cleaned up by HBase, leading to excess storage usage and possible error due to filling up the storage. Accumulated WAL files can lead to lengthy restart times as they must all be played back to ensure no dataloss occurs on restart. This can have follow-on HDFS impact if the number of WAL files overwhelm HDFS Name Node.

Upstream JIRA

- [HBASE-20781](#)
- [HBASE-25459](#)
- [PHOENIX-5250](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-494: Accumulated WAL Files Cannot be Cleaned up When Using Phoenix Secondary Global Indexes](#)

TSB 2021-506: Active HBase MOB files can be removed

Actively used MOB files can be deleted by MobFileCleanerChore due to incorrect serialization of reference file names. This is causing data loss on MOB-enabled tables.

Upstream JIRA

- [HBASE-23723](#)
- [HBASE-25970](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-506: Active HBase MOB files can be removed](#)

Known Issues in HDFS

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

OPSAPS-60958: The `dfs.access.time.precision` and `dfs.namenode.access.time.precision` parameters are available in Cloudera Manager > HDFS > Configuration.

You must configure both the `dfs.access.time.precision` and `dfs.namenode.access.time.precision` parameters with the same value as Cloudera Manager still sends both the parameters to HDFS service configuration.

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 2021-458: Possible HDFS Erasure Coded (EC) Data Files Corruption in EC Reconstruction

Cloudera has detected two bugs that can cause corruption of HDFS Erasure Coded (EC) files during the data reconstruction process.

The first bug can be hit during DataNode decommissioning. Due to a bug in the data reconstruction logic during decommissioning, some parity blocks may be generated with a content of all zeros.

The second issue occurs in a corner case when a DataNode times out in the reconstruction process. It will reschedule a read from another good DataNode. However, the stale DataNode reader may have polluted the buffer and subsequent reconstruction which uses the polluted buffer will suffer from EC block corruption.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [Cloudera Customer Advisory: Possible HDFS Erasure Coded \(EC\) Data Files Corruption in EC Reconstruction](#)

Known Issues in Apache Hive

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

OPSAPS-58664: Hive on Tez LDAP configurations are not pushed to hive-site.xml by Cloudera Manager

After setting up LDAP properties in the Hive on Tez service, the settings are not pushed into hive-site.xml for Hive on Tez service even after a restart. The issue is due to HiveOnTezServiceHandler re-using definitions from HiveConfigFileDefinitions. The definitions are not including any roletypes other than HiveServiceHandler's roletypes.

OPSAPS-59928: INSERT INTO from SELECT using hive (hbase) table returns an error under certain conditions.

Users who upgraded to a Kerberized CDP cluster from HDP and enabled AutoTLS have reported this problem. For more information, see [Cloudera Community article: ERROR: "FAILED: Execution Error, return code 2" when the user is unable to issue INSERT INTO from SELECT using hive \(hbase\) table](#).

In Cloudera Manager TEZ Configurations, find the tez.cluster.additional.classpath.prefix Safety Valve, and set the value to /etc/hbase/conf.

CDPD-17507: After upgrading from HDP to Cloudera Runtime 7.1.3 tables created with Spark SQL write to wrong HDFS location.

Remove spark.sql.warehouse.dir from spark-defaults or upgrade to Cloudera Runtime 7.1.4.

CDPD-21365: Performing a drop catalog operation drops the catalog from the CTLGS table. The DBS table has a foreign key reference on CTLGS for CTLG_NAME. Because of this, the DBS table is locked and creates a deadlock.

You must create an index in the DBS table on CTLG_NAME: CREATE INDEX CTLG_NAME_DBS ON DBS(CTLG_NAME);.

OPSAPS-60546: Upgrading from CDH to Cloudera Runtime 7, the Hive Java Heap Size does not propagate and defaults to 2GB.

Manually reconfigure Hive Java Heap Size after upgrade.

OPSAPS-54299 Installing Hive on Tez and HMS in the incorrect order causes HiveServer failure

You need to install Hive on Tez and HMS in the correct order; otherwise, HiveServer fails. You need to install additional HiveServer roles to Hive on Tez, not the Hive service; otherwise, HiveServer fails.

Workaround: Follow instructions on [Installing Hive on Tez](#).

CDPD-23041: DROP TABLE on a table having an index does not work

If you migrate a Hive table to CDP having an index, DROP TABLE does not drop the table. Hive no longer supports indexes ([HIVE-18448](#)). A foreign key constraint on the indexed table prevents dropping the table. Attempting to drop such a table results in the following error:

```
java.sql.BatchUpdateException: Cannot delete or update a parent
row: a foreign key constraint fails ("hive"."IDXS", CONSTRAINT "
IDXS_FK1" FOREIGN KEY ("ORIG_TBL_ID") REFERENCES "TBL" ("TBL_ID"
))
```

There are two workarounds:

- Drop the foreign key "IDXS_FK1" on the "IDXS" table within the metastore. You can also manually drop indexes, but do not cascade any drops because the IDXS table includes references to "TBL".
- Launch an older version of Hive, such as Hive 2.3 that includes IDXS in the DDL, and then drop the indexes as described in [Language Manual Indexing](#).

Apache Issue: [Hive-24815](#)

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`)

CDPD-16802: Autotranslate assertion failure.

The exception is not triggered when it is executed from Spark-Shell. This is from Hive in the `getJdoFilterPushdownParam` parameter of `ExpressionTree.java`, which checks the partition column as only String and not any other type.

This can be disabled by setting `hive.metastore.integral.jdo.pushdown` to true.

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, CDPD-10352: 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`.

HIVE-24271: Problem creating an ACID table in legacy table mode

If you configured legacy CREATE TABLE behavior at the site level, the CREATE MANAGED TABLE command might not create a managed ACID table. This command works only at the session level.

Workaround: Include table properties in a CREATE TABLE that specify a transactional table. For example:

```
CREATE TABLE T2(a int, b int)
STORED AS ORC
TBLPROPERTIES ('transactional'='true');
```

Technical Service Bulletins

TSB 2021-459: Renaming managed (ACID) table shows empty records

Renaming an ACID (managed) table using ALTER TABLE <table name> RENAME causes empty records in the table. Also, the location of the new table after renaming points to the location of the old table before renaming. This can cause correctness issues, for example:

```
create table abc (id int);
insert into abc values (1);
rename table abc to def; create table abc (id int); // should be
empty
insert into abc values (2);
select * from abc ; // returns 1 and 2, the new and the old re
sults
```

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-459: Renaming managed \(ACID\) table shows empty records](#)

TSB 2021-480/1: Hive produces incorrect query results when skipping a header in a binary file

In CDP, setting the table property skip.header.line.count to greater than 0 in a table stored in a binary format, such as Parquet, can cause incorrect query results. The skip header property is intended for use with Text files and typically used with CSV files. The issue is not present when you run the query on a Text file that sets the skip header property to 1 or greater.

Upstream JIRA

[Apache Jira: HIVE-24827](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB-2021 480.1: Hive produces incorrect query results when skipping a header in a binary file](#)

TSB 2021-480/2: Hive ignores the property to skip a header or footer in a compressed file

In CDP, setting the table properties skip.header.line.count and skip.footer.line.count to greater than 0 in a table stored in a compressed format, such as bzip2, can cause incorrect results from SELECT * or SELECT COUNT (*) queries.

Upstream JIRA

[Apache Jira: HIVE-24224](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-480.2: Hive ignores the property to skip a header or footer in a compressed file](#)

TSB 2021-482: Race condition in subdirectory delete/rename causes hive jobs to fail

Multiple threads try to perform a rename operation on s3. One of the threads fails to perform a rename operation, causing an error. Hive logs will report "HiveException: Error moving ..." and the log will contain an error line starting with " Exception when loading partition " -all paths listed with s3a:// prefixes.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-482: Race condition in subdirectory delete/rename causes Hive jobs to fail](#)

TSB 2021-501: JOIN queries return wrong result for join keys with large size in Hive

JOIN queries return wrong results when performing joins on large size keys (larger than 255 bytes). This happens when the fast hash table join algorithm is enabled, which is enabled by default.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-501: JOIN queries return wrong result for join keys with large size in Hive](#)

TSB 2021-518: Incorrect results returned when joining two tables with different bucketing versions

Incorrect results are returned when joining two tables with different bucketing versions, and with the following Hive configurations: set hive.auto.convert.join = false and set mapreduce.job.reduces = any custom value.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-518: Incorrect results returned when joining two tables with different bucketing versions](#)

TSB 2021-524: Intermittent data duplication if direct insert enabled

If direct insert is enabled, data is written directly to the final location with an attemptId. At the end of the insert operation, all data written before the final attempt should be deleted. However due to a bug in HIVE-21164, this does not happen.

Example: Data is written to the final location with attemptId=0, but this task fails. Hive tries the task again and writes data to the final location with attemptId=1. At the end of the insert, Hive should remove all the files with attemptId=0, but it does not.

Upstream JIRA

- [HIVE-21164](#)
- [HIVE-24322](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-524: Intermittent data duplication if direct insert enabled](#)

TSB 2022-526: A Hive query may produce wrong results for some vectorized built-in functions with compound expression in PARTITION BY or ORDER BY clause

Vectorized functions with PARTITION BY and/or ORDER BY clauses where the partition or order by expression is compound (example: cast string to integer) and not just a simple column reference may be broken.

The query may fail or output wrong results, depending on the compound expression. For example:

- Cast integer to string results in query failure with a NullPointerException
- Cast string to integer outputs wrong results

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-526: A Hive query may produce wrong results for some vectorized built-in functions with compound expression in PARTITION BY or ORDER BY clause](#)

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.

CDPD-40354: Uploading large files to WebHDFS from Hue secured using Knox fails

When you upload files larger than 300 MB to WebHDFS using Hue File Browser from a Hue instance that is secured using Knox, the upload fails and you see the following error on the Hue web interface: 502 bad gateway nginx error. This error occurs only when you access the Hue instance from the Knox Gateway. Also, you may not find any exceptions in Hue or Knox service logs.

This issue is caused because the value of the “addExpect100Continue” is set to “true”. Therefore streaming large files causes out of memory error in Knox.

To resolve the issue, you must set the value of the “addExpect100Continue” to “false”.

1. SSH in to the Knox host as an Administrator.
2. Back up the following files:
 - /opt/cloudera/parcels/CDH-**[***VERSION***]**/lib/knox/data/services/hue/1.0.0/service.xml
 - /var/lib/knox/gateway/data/services/hue/1.0.0/service.xml
3. Open the /opt/cloudera/parcels/CDH-**[***VERSION***]**/lib/knox/data/services/hue/1.0.0/service.xml file for editing, search for the addExpect100Continue property, and set the value of this property to “false” as follows:

```
<param>
  <name>addExpect100Continue</name>
  <value>>false</value>
</param>
```

4. Save the file and exit.
5. Open the /var/lib/knox/gateway/data/services/hue/1.0.0/service.xml file for editing, search for the addExpect100Continue property, and set the value of this property to “false” as follows:

```
<param>
  <name>addExpect100Continue</name>
  <value>>false</value>
</param>
```

6. Save the file and exit.
7. Delete the deployment directory as follows:

```
rm -rf /var/lib/knox/gateway/data/deployments/
```

8. Restart the Knox service.

Hue uses the unsafe-inline directive in its Content Security Policy (CSP) header

Hue 4 web interface uses the unsafe-inline directive in its CSP header. As a result, the application server does not set the CSP header in its HTTP responses, and therefore does not benefit from the additional protection against potential cross-site scripting issues and other modern application vulnerabilities which a properly configured CSP may provide. This could lead to application vulnerability.

This issue will be fixed in Hue 5. Until then, Cloudera recommends deploying additional security measures such as a firewall within the Hue server to control allowed connections, and SSO-based authentications mechanisms such as LDAP or SAML.

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_original
```

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 queuing 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.

Psycopg2 library needed for PostgreSQL-backed Hue when on RHEL 8 or Ubuntu 20 platforms

You may see a warning on the **Host Inspector Results** page stating that a compatible version of the Psycopg2 library is missing on your host if you have installed CDP 7.1.7 on RHEL 8 or Ubuntu 20 platforms and if you are using PostgreSQL as the backend database for Hue. This is because RHEL 8 and Ubuntu 20 contain Python 3 by default and Hue does not support Python 3.

Hue in Runtime 7 requires version 2.7.5 of the psycopg2 Python package for connecting to a PostgreSQL database. If you are on RHEL 8 or Ubuntu 20, then you must install one of the following compatible psycopg2 package versions before deploying CDP 7.1.7 on your cluster: 2.7.5, 2.7.6.1, and 2.7.7. For more information, see [Installing the psycopg2 Python Package on PostgreSQL-backed Hue](#).

Impala editor fails silently after SAML SSO session times out

When you run a query from an Impala editor in Hue, the Impala editor may silently fail without displaying an error message. As a result, you may not see any action on the screen after submitting your query. This happens if Hue is configured with SAML authentication and you run a query from a browser session that has remained open for a period longer than the SSO maximum session time set by the SAML Identity Provider.

If you do not see any action on the Impala editor after submitting the query, refresh the page on your browser.

Connection failed error when accessing the Search app (Solr) from Hue

If you are using Solr with Hue to generate interactive dashboards and for indexing data, and if you have deployed two Solr services on your cluster and selected the second one as a dependency for Hue, then Cloudera Manager assigns the hostname of the first Solr service and the port number of the second Solr service generating an incorrect Solr URL in the search section of the hue.ini file. As a result, you may see a “Connection failed” error when you try to access the Search app from the Hue web UI.

1. Log into Cloudera Manager as an Administrator.
2. Go to Clusters Hue service Configuration and add the following lines in the Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini field:

```
[search]
# URL of the Solr Server
solr_url=http://[***HOSTNAME***]:[***PORT***]/solr/
```

For example:

```
solr_url=http://solr2:4567/solr/
```

3. Click Save Changes.
4. Restart the Hue service.

Invalid S3 URI error while accessing S3 bucket

The Hue Load Balancer merges the double slashes (//) in the S3 URI into a single slash (/) so that the URI prefix "/filebrowser/view=S3A://" is changed to "/filebrowser/view=S3A:". This results in an error when you try to access the S3 buckets from the Hue File Browser through the port 8889.

The Hue web UI displays the following error: “Unknown error occurred”.

The Hue server logs record the “ValueError: Invalid S3 URI: S3A” error.

To resolve this issue, add the following property in the Hue Load Balancer Advanced Configuration Snippet:

1. Sign in to Cloudera Manager as an administrator.
2. Go to Clusters Hue service Configurations Load Balancer and search for the Load Balancer Advanced Configuration Snippet (Safety Valve) for httpd.conf field.
3. Specify MergeSlashes OFF in the Load Balancer Advanced Configuration Snippet (Safety Valve) for httpd.conf field.
4. Click Save Changes.
5. Restart the Hue Load Balancer.

You should be able to load the S3 browser from both 8888 and 8889 ports.

Alternatively, you can use the Hue server port 8888 instead of the load balancer port 8889 to resolve this issue.

Error while rerunning Oozie workflow

You may see an error such as the following while rerunning an already executed and finished Oozie workflow through the Hue web interface: E0504: App directory [hdfs://cdh/user/hue/oozie/workspaces/hue-oozie-1571929263.84] does not exist.

To resolve this issue, add the following property in the Hue Load Balancer Advanced Configuration Snippet:

1. Sign in to Cloudera Manager as an administrator.

2. Go to Clusters Hue service Configurations Load Balancer and search for the Load Balancer Advanced Configuration Snippet (Safety Valve) for httpd.conf field.
3. Specify MergeSlashes OFF in the Load Balancer Advanced Configuration Snippet (Safety Valve) for httpd.conf field.
4. Click Save Changes.
5. Restart the Hue Load Balancer.

Python-psycopg2 package version 2.8.4 not compatible with Hue

Ubuntu 18.04 provides python-psycopg2 package version 2.8.4 but it is not compatible with Hue because of a bug in the Django framework.

Downgrade the package at the OS level by running the following command:

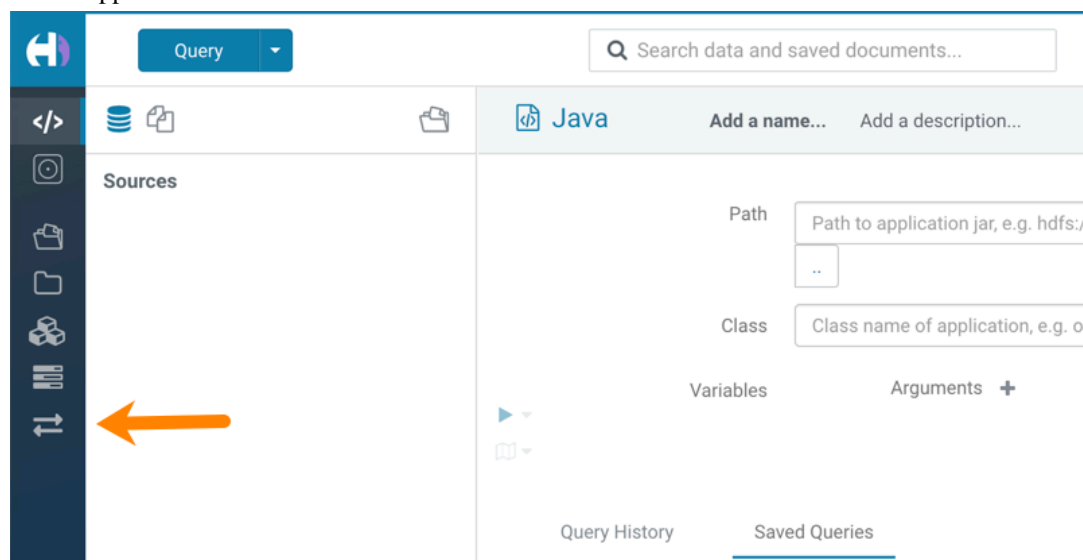
```
sudo apt install python-psycopg2==2.7.5
```

or install python-psycopg2 package using pip by running the following command:

```
sudo pip install psycopg2==2.7.5
```

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:



Hue limitation after upgrading from CDH to CDP Private Cloud Base

The `hive.server2.parallel.ops.in.session` configuration property changes from TRUE to FALSE after upgrading from CDH to CDP Private Cloud Base. Current versions of Hue are compatible with this property change; however, if you still would like to use an earlier version of Hue that was not compatible with this property being FALSE and shared a single JDBC connection to issue queries concurrently, the connection will no longer work after upgrading.

Unsupported feature: 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.

PySpark and SparkSQL are not supported with Livy in Hue

Hue does not support configuring and using PySpark and SparkSQL with Livy in CDP Private Cloud Base.

Technical Service Bulletins

TSB 2021-487: Cloudera Hue is vulnerable to Cross-Site Scripting attacks

Multiple Cross-Site Scripting (XSS) vulnerabilities of Cloudera Hue have been found. They allow JavaScript code injection and execution in the application context.

- CVE-2021-29994 - The Add Description field in the Table schema browser does not sanitize user inputs as expected.
- CVE-2021-32480 - Default Home direct button in Filebrowser is also susceptible to XSS attack.
- CVE-2021-32481 - The Error snippet dialog of the Hue UI does not sanitize user inputs.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-487: Cloudera Hue is vulnerable to Cross-Site Scripting attacks \(CVE-2021-29994, CVE-2021-32480, CVE-2021-32481\)](#)

Known Issues in Apache Impala

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

IMPALA-2636: HS2 GetTables() returns TABLE_TYPE as TABLE for VIEW

Impala returns the TABLE_TYPE value as "Table" when querying tables using HS2's GetTables endpoint. Due to this, the views are identified (visible) as tables in Impala's Hue editor (you would see a table icon in front of the view rather than a view icon), though the actual working remains the same as views.

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 HDFS file hdfs://nameservice1/warehouse/tablespace/managed/hive/<database>/<table>/xxxxx
Error(2): No such file or directory Root cause: RemoteException: File does not exist: /warehouse/tablespace/managed/hive/<database>/<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
```

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-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.

None

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.

None

IMPALA-2603: Crash: impala::Coordinator::ValidateCollectionSlots

A query could encounter a serious error if includes multiple nested levels of INNER JOIN clauses involving subqueries.

None

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
```

[illegible]

IMPALA-3509: Breakpad minidumps can be very large when the thread count is high

The size of the breakpoint 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 RECOVER PARTITIONS`, may delay metadata propagation of unrelated unloaded tables triggered by statements like `DESCRIBE` or `SELECT` queries.

None

IMPALA-7072: Impala does not support Heimdal Kerberos

None

OPSAPS-46641: A single parameter exists in Cloudera Manager for specifying the Impala Daemon Load Balancer. Because BDR and Hue need to use different ports when connecting to the load balancer, it is not possible to configure the load balancer value so that BDR and Hue will work correctly in the same cluster.

The workaround is to use the load balancer configuration either without a port specification, or with the Beeswax port: this will configure BDR. To configure Hue use the "Hue Server Advanced Configuration Snippet (Safety Valve) for impalad_flags" to specify the the load balancer address with the HiveServer2 port.

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.

Some of the unresolved issues include:

- [IMPALA-6841](#)
- [HADOOP-15720](#)
- [IMPALA-635](#)

Technical Service Bulletins

TSB-2021-485: Impala returns fewer rows from parquet tables on S3

[IMPALA-10310](#) was an issue in Impala's Parquet page filtering code where the scanner did not reset state appropriately when transitioning from the first row group to subsequent row groups in a single split. This caused data from the subsequent row groups to be skipped incorrectly, leading to incorrect query results. This issue cannot occur when the Parquet page filtering is disabled by setting `PARQUET_READ_PAGE_INDEX=false`.

The issue is more likely to be encountered on S3/ADLS/ABFS/etc, because Spark is sometimes configured to write 128MB row groups and the `PARQUET_OBJECT_STORE_SPLIT_SIZE` is 256MB. This makes it more likely for Impala to process two row groups in a single split.

Parquet page filtering only works based on the min/max statistics, therefore the comparison operators it supports are “=”, “<”, “>”, “<=”, and “>=”. These operators are impacted by this bug. Expressions such as “!=”, 'LIKE' or the expressions including UDF do not use parquet page filtering.

The `PARQUET_OBJECT_STORE_SPLIT_SIZE` parameter is introduced in Impala 3.3 by [IMPALA-5843](#). This means that older versions of Impala do not have this issue.

Upstream JIRA

- [IMPALA-5843](#)
- [IMPALA-10310](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB-2021-485: Impala returns fewer rows from parquet tables on S3](#)

TSB 2021-502: Impala logs the session / operation secret on most RPCs at INFO level

Impala logs contain the session / operation secret. With this information a person who has access to the Impala logs might be able to hijack other users' sessions. This means the attacker is able to execute statements for which they do not have the necessary privileges otherwise. Impala deployments where Apache Sentry or Apache Ranger authorization is enabled may be vulnerable to privilege escalation. Impala deployments where audit logging is enabled may be vulnerable to incorrect audit logging.

Restricting access to the Impala logs that expose secrets will reduce the risk of an attack. Additionally, restricting access to trusted users for the Impala deployment will also reduce the risk of an attack. Log redaction techniques can be used to redact secrets from the logs. For more information, see the *Cloudera Manager documentation*.

For log redaction, users can create a rule with a search pattern: `secret \((string\) [=:].*` And the replacement could be for example: `secret=LOG-REDACTED`

Upstream JIRA

[IMPALA-10600](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-502: Impala logs the session / operation secret on most RPCs at INFO level](#)

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](#)

TSB 2023-632: Apache Impala reads minor compacted tables incorrectly on CDP Private Cloud Base

The issue occurs when Apache Impala (Impala) reads insert-only Hive ACID tables that were minor compacted by Apache Hive (Hive).

Insert-only ACID table (also known as micro-managed ACID table) is the default table format in Impala in CDP Private Cloud Base 7.1.x and can be identified by having the following table properties:

```
"transactional"="true"
"transactional_properties"="insert_only"
```

Minor compactions can be initiated in Hive with the following statement:

```
ALTER TABLE <table_name> COMPACT 'minor'
```

A minor compaction differs from a major compaction in compacting only the files created by INSERTs since the last compaction instead of compacting all files in the table.

Performing a minor compaction results in creation of delta directories in the table (or partition) folder like `delta_0000001_0000008_v0000564`. These delta directories are not handled correctly by Impala, which can lead to returning different results compared to Hive. This means either missing rows from some data files or duplicating rows from some data files. The exact results depend on whether a major compaction was run on the table and on whether the old files compacted during a minor compaction have been deleted.

If the last compaction was a major compaction or if neither a minor nor a major compaction was performed on the table, then the issue does not occur.

Minor compaction is not initiated automatically by Hive Metastore (HMS) or any other CDP (Cloudera Data Platform) component, meaning that this issue can only occur if minor compactions were initiated explicitly by users or scripts.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-632 Impala reads minor compacted tables incorrectly on CDP Private Cloud Base](#)

Known Issues in Apache Kafka

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

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-configs
- 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.

Custom Kerberos principal names cannot be used for kerberized ZooKeeper and Kafka instances

When using ZooKeeper authentication and a custom Kerberos principal, Kerberos-enabled Kafka does not start. You must disable ZooKeeper authentication for Kafka or use the default Kerberos principals for ZooKeeper and Kafka.

None

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-59031: Kafka cannot start if configuration is added to the Kafka Broker Advanced Configuration Snippet (Safety Valve) for `ssl.properties`

The Kafka Broker Advanced Configuration Snippet (Safety Valve) for `ssl.properties` configuration snippet does not correctly override configuration. As a result, Kafka may not start if TLS/SSL related configuration overrides are added to the this configuration snippet.

Use the Kafka Broker Advanced Configuration Snippet (Safety Valve) for `kafka.properties` configuration snippet instead to override SSL related properties.

OPSAPS-58492: Automatic Ranger service creation for Kafka is incomplete

When a new resource based service is created for Kafka in Ranger, the `cruisecontrol` user is not added to the policy set. This prevents Cruise Control from working properly.

Manually add the `cruisecontrol` user to the Kafka resource based service in Ranger.

CDPD-11775: Kafka Connect does not start due to occupied ports

By default the Kafka Connect role binds to ports that are in the ephemeral range. As many other services can use port 0 and can bind to any port in the ephemeral range, it can happen that the default Kafka Connect ports become occupied. In a case like this, the Kafka Connect role will not start.

Configure the Kafka Connect rest port, Kafka Connect secure rest port, and Jetty Metrics port to expose JMX Json Kafka properties in Cloudera Manager. Cloudera recommends that you use the following ports:

- Kafka Connect rest port: 28083
- Kafka Connect secure rest port: 28085
- Jetty Metrics port to expose JMX Json: 28084

Additionally, if you are using SMM to monitor Kafka Connect, you must also configure the Kafka Connect Rest PortStreams Messaging Manager property. The port configured in this property must match the port configured for the Kafka Connect role.

Unsupported Features

The following Kafka features are not supported in Cloudera Data Platform:

- Only Java and .Net based clients are supported. Clients developed with C, C++, Python, and other languages are currently not supported.
- The Kafka default authorizer is not supported. This includes setting ACLs and all related APIs, broker functionality, and command-line tools.

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 Hosts 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_entity_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 Kerberos

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

OPSAPS-60331: If Cloudera Manager is configured to use Active Directory as a Kerberos KDC, and is also configured to use /etc/cloudera-scm-server/cmf.keytab as the KDC admin credentials, you may encounter errors when generating Kerberos credentials.

In the Cloudera Manager Admin Console, run the "Administration > Security > Kerberos Credentials > Import KDC Account Manager Credentials" wizard. Remove /etc/cloudera-scm-server/cmf.keytab on the Cloudera Manager server host.

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.

OPSAPS-58179: HIVE endpoint url is updated on only one Knox host topology. While on other Knox host, the Cloudera Manager configuration monitoring change is not identified and topologies are not updated with the Hive URL.

None

OPSAPS-59751: If Cloudera Manager is configured with Apache Knox, then Replication Manager does not work.

None

Technical Service Bulletins

TSB 2022-553: DOM based XSS Vulnerability in Apache Knox

When using Knox Single Sign On (SSO) in the affected releases, a request could be crafted to redirect a user to a malicious page due to improper URL parsing. The request includes a specially crafted request parameter that could be used to redirect the user to a page controlled by an attacker. This request URL would need to be presented to the user outside the normal request flow through a XSS or phishing campaign.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-553: DOM based XSS Vulnerability in Apache Knox \(“Knox”\)](#)

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.
- Kudu HMS Sync is disabled and is not yet supported

Known Issues in Apache Oozie

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

OOZIE-3549: Oozie fails to start when Cloudera Manager 7.x is used with Cloudera Runtime 6.x and Java 11 because Oozie does not set the trust-store password.

The issue is fixed in OOZIE-3549 and is already included in CDP 7.x but not in CDH 6.x. If you are on CDH 6.x and want to upgrade to Java 11 or your Cloudera Manager to 7.x then you must request for a patch.

OPSAPS-57429: Zookeeper SSL or TLS support for Oozie.

When SSL is enabled in Zookeeper, Oozie tries to connect to Zookeeper using SSL instead of a non-secure connection.

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.

Known Issues in Ozone

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

CDPD-15268:Uploading a key using the S3 Multi-part upload API into an Ozone encryption zone (TDE-enabled bucket) is not currently supported. The key upload will fail with an exception.

None

CDPD-15362:When files and directories stored in Ozone are deleted via Hadoop filesystem shell `-rm` command using `o3fs` or `ofs` scheme, they will not be moved to trash even if `fs.trash.interval` is set to `>0` on the client. Instead, they are deleted immediately.

None

CDPD-15330:

A network partitioned Ozone Manager (OM) in an OM HA cluster, if in a leader state before partition and does not step down as leader, can serve stale reads.

No workaround is available. If network partition is detected and the Ozone Manager node is restarted, then this issue can be resolved (even if network partition exists after restart).

CDPD-15266:When Ozone Manager (OM) HA is enabled, not all older OM Ratis logs are purged. Similarly, for DataNode, old Ratis logs may not be purged. This can lead to older logs consuming the disk space.

For OM, you must manually delete the OM Ratis logs from the Ratis storage directory location defined by the `ozone.om.ratis.storage.dir` property. You must only delete the logs older than the already purged logs.

For example, if the OM Ratis log directory contains the logs `log_0_100`, `log_101_200`, and `log_301_400`, then you can delete `log_0_100` and `log_101_200` as `log_201_300` is already purged.

For DataNode, you must manually delete `datanodeRatis` logs from the Ratis storage directory location defined by the `dfs.container.ratis.datanode.storage.dir` property. You must delete only the logs older than already purged logs.

For example, if the DataNode Ratis log directory contains the logs `log_0_100`, `log_101_200`, and `log_301_400`, then you can delete `log_0_100` and `log_101_200` as `log_201_300` is already purged.

Cloudera advises you to backup the ratis logs. Ensure that DataNode and OM come as back up again and the pipelines they are connected to must be healthy. In case there are any exceptions, the Ratis logs must be restored from the backup.

CDPD-15602: Creating or deleting keys with a trailing forward slash (/) in the name is not supported via the Ozone shell or the S3 REST API. Such keys are internally treated as directories by the Ozone service for compatibility with the Hadoop filesystem interface. This will be supported in a later release of CDP.

You can create or delete keys via the Hadoop Filesystem interface, either programmatically or via the filesystem Hadoop shell. For example, ``ozone fs -rmdir <dir>``.

Technical Service Bulletins

TSB 2021-523: Multiple CVEs - Ozone security identified and addressed

The following CVEs have been addressed:

CVE link	CVE title	Affects versions
CVE-2021-36372	Original block tokens are persisted and can be retrieved	7.1.3 until 7.1.6
CVE-2021-39231	Missing authentication/authorization on internal RPC endpoints	7.1.3 until 7.1.6
CVE-2021-39232	Missing admin check for SCM related admin commands	7.1.3 until 7.1.5
CVE-2021-39233	Container-related datanode operations can be called without authorization	7.1.3 until 7.1.6

CVE link	CVE title	Affects versions
CVE-2021-39234	Raw block data can be read bypassing ACL/ authorization	7.1.3 until 7.1.6
CVE-2021-39235	Access mode of block tokens are not enforced	7.1.3 until 7.1.5
CVE-2021-39236	Owners of the S3 tokens are not validated	7.1.3 until 7.1.5
CVE-2021-41532	Unauthenticated access to Ozone Recon HTTP endpoints	7.1.3 until 7.1.7

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-523: Multiple CVEs - Ozone security identified and addressed](#)

Known Issues in Apache Phoenix

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

Technical Service Bulletins

TSB 2022-568: HBase normalizer must be disabled for Salted Phoenix tables

When Apache Phoenix (“Phoenix”) creates a salted table, it pre-splits the table according to the number of salt regions. These regions must always be kept separate, otherwise Phoenix does not work correctly.

The HBase normalizer is not aware of this requirement, and in some cases the pre-split regions are merged automatically. This causes failure in Phoenix.

The same requirement applies when merging regions of salted tables manually: regions containing different salt keys (the first byte of the rowkey) must never be merged.

Note that either automatic or manual splitting of the regions for a salted table does not cause a problem. The problem only occurs when adjacent regions containing different salt keys are merged.

Upstream JIRA

[PHOENIX-4906](#)

Knowledge article

For the latest update on this issue, see the corresponding Knowledge article: [TSB 2022-568: Hbase normalizer must be disabled for Salted Phoenix tables](#)

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.

Known Issues in Schema Registry

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

CDPD-18345: Schema Registry fails to start on a cluster with TLS enabled and multiple SANs in the certificate

Schema Registry fails to start on a cluster if TLS is enabled and the certificate contains multiple Subject Alternative Names (SAN).

Use a certificate that does not have multiple SANs. Or, temporarily disable TLS for the Schema Registry service to start.

DOCS-7429: Schema Registry Ranger Policy requires schemaregistry group in Ranger Admin

Schema Registry fails to start because the schemaregistry user group does not exist in Ranger. The following message appears: Operation denied. Group name: schemaregistry specified in policy does not exist in ranger admin.

Ensure that the schemaregistry user group exists in Ranger. Depending on the environment settings, this might require manual creation of users/groups and/or re-configuration of the usersync/groupsync filter in Ranger.

Known Issues in Cloudera Search

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

Splitshard of 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.

None

DOCS-5717: Lucene index handling limitation

The Lucene index can only be upgraded by one major version. Solr 8 will not open an index that was created with Solr 6 or earlier.

None, you need to reindex collections.

CDH-82042: 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 collection --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.

CCDPD-4139: Collection state goes down after Solr SSL.

If you enable TLS/SSL on a Solr instance with existing collections, the collections will break and become unavailable. Collections created after enabling TLS/SSL are not affected by this issue.

Recreate the collection after enabling TLS.

CCPD-13923: Every Configset is Untrusted Without Kerberos

Solr 8 introduces the concept of ‘[untrusted configset](#)’, denoting configsets that were uploaded without authentication. Collections created with an untrusted configset will not initialize if `<lib>` directives are used in the configset.

Select one of the following options if you would like to use untrusted configsets with `<lib>` directives:

- If the configset contains external libraries, but you do not want to use them, simply upload the configsets after deleting the `<lib>` directives.
- If the configset contains external libraries, and you want to use them, choose one from the following options:
 - Secure your cluster before reuploading the configset.
 - Add the libraries to Solr’s classpath, then reupload the configset without the `<lib>` directives.

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).

Limitations

Default Solr core names cannot be changed

Although it is technically possible to give user-defined Solr core names during core creation, it is to be avoided in the context of Cloudera Search. Cloudera Manager expects core names in the default "collection_shardX_replicaY" format. Altering core names results in Cloudera Manager being unable to fetch Solr metrics for the given core and this, eventually, may corrupt data collection for co-located core, or even shard and server level charts.

Known Issues in Apache Solr

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

Technical Service Bulletins

TSB 2021-495: CVE-2021-29943: Apache Solr Unprivileged users may be able to perform unauthorized read/write to collections

Using the ConfigurableInternodeAuthHadoopPlugin class as the authentication plugin with Ranger as the authorization module introduced a backdoor for unauthorized access to data. With this combination, when an authenticated user sends a query to a node, which does not have the data locally, the request will be forwarded in the name of the Solr service user and not in the name of the original requester. In this case, the authorization happens against the user named solr which may have almost full access. It may be the case that infra Solr customers were advised to switch back to ConfigurableInternodeAuthHadoopPlugin. Only these customers should be affected by this CVE.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-495: Apache Solr Unprivileged users may be able to perform unauthorized read/write to collections - CVE-2021-29943](#)

TSB 2021-497: CVE-2021-27905: Apache Solr SSRF vulnerability with the Replication handler

The Apache Solr ReplicationHandler (normally registered at "/replication" under a Solr core) has a "masterUrl" (also "leaderUrl" alias) parameter. The "masterUrl" parameter is used to designate another ReplicationHandler on another Solr core to replicate index data into the local core. To help prevent the CVE-2021-27905 SSRF vulnerability, Solr should check these parameters against a similar configuration used for the "shards" parameter.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-497: CVE-2021-27905: Apache Solr SSRF vulnerability with the Replication handler](#)

Known Issues in Apache Spark

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

CDPD-22670 and CDPD-23103: There are two configurations in Spark, "Atlas dependency" and "spark_lineage_enabled", which are conflicted. The issue is when Atlas dependency is turned off but spark_lineage_enabled is turned on.

Run Spark application, Spark will log some error message and cannot continue. That can be restored by correcting the configurations and restarting Spark component with distributing client configurations.

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.vectorized.use.checked.expressions does not exist
19/08/09 11:48:04 WARN conf.HiveConf: HiveConf of name hive.tez.cartesian-product.enabled does not exist
```

These errors can be safely ignored.

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

MM2-163: 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-14019: SRM may automatically re-create deleted topics

If `auto.create.topics.enable` is enabled, deleted topics are automatically recreated on source clusters.

Prior to deletion, remove the topic from the topic allowlist with the `srm-control` tool. This prevents topics from being re-created.

```
srm-control topics --source [SOURCE_CLUSTER] --target [TARGET_CLUSTER] --remove [TOPIC1][TOPIC2]
```

CSP-462: Replication failing when SRM driver is present on multiple nodes

Kafka replication fails when the SRM driver is installed on more than one node.

None

CDPD-11074: The `srm-control` tool can be called without `--target`

The `srm-control` tool can be initialized without specifying the `--target` option. If the tool is called this way it will fail to run correctly.

Do not use the tool without specifying the `--target` option. Always specify both `--source` and `--target` options. For example:

```
srm-control topics --source [SOURCE_CLUSTER] --target [TARGET_CLUSTER] --list
```

CDPD-13864 and CDPD-15327: Replication stops after the network configuration of a source or target cluster is changed

If the network configuration of a cluster which is taking part in a replication is changed, for example, port numbers are changed as a result of enabling or disabling TLS, SRM will not update its internal configuration even if SRM is reconfigured and restarted. From SRM's perspective, it is the cluster identity that has changed. SRM cannot determine whether the new identity corresponds to the same cluster or not, only the owner or administrator of that cluster can know. In this case,

SRM tries to use the last known configuration of that cluster which might not be valid, resulting in the halt of replication.

There are three workarounds for this issue. Choose one of the following:

Increase the driver rebalance timeout

Increasing the rebalance timeout to 5 minutes (300000 ms) or longer can resolve the issue. In general a 5 minute timeout should be sufficient for most deployments. However, depending on your scenario, an even longer period might be required. Increasing the rebalance timeout might lead to increased latency when the SRM drivers stop. The cluster will be slower when it rebalances the load of the removed driver.

The rebalance timeout can be configured on a per cluster (alias) basis by adding the following to the Streams Replication Manager's Replication Configs Cloudera Manager property:

```
[***ALIAS***].rebalance.timeout.ms = [***VALUE***]
```

Replace [***ALIAS***] with a cluster alias specified in Streams Replication Manager Cluster alias. Do this for all clusters that are taking part in the replication process. When correctly configured, your configuration will have a rebalance.timeout.ms entry corresponding to each cluster (alias). For example:

```
primary.rebalance.timeout.ms = 30000
secondary.rebalance.timeout.ms = 30000
tertiary.rebalance.timeout.ms = 30000
```

After the new broker configuration is applied by SRM, the rebalance timeout can be reverted back to its original value, or removed from the configuration altogether.

Decrease replication admin timeout

Decreasing the replication admin timeout to 15 seconds (15000 ms) can resolve the issue. With higher loads, this might cause WARN messages to appear in the SRM driver log.

The admin timeout can be configured on a per replication basis by adding the following to the Streams Replication Manager's Replication Configs Cloudera Manager property:

```
[***REPLICATION***].admin.timeout.ms = [***VALUE***]
```

Replace [***REPLICATION***] with a replication specified in Streams Replication Manager's Replication Configs. Do this for all affected replications. When correctly configured, your configuration will have an admin.timeout.ms entry corresponding to each affected replication. For example:

```
primary->secondary.admin.timeout.ms = 15000
secondary->primary.admin.timeout.ms = 15000
```

After the new broker configuration is applied by SRM, the admin timeout can be reverted back to its original value, or removed from the configuration altogether.

Upgrade the brokers incrementally

Instead of switching over to the new configuration, open two separate listeners on the broker. One for the old configuration, and one for the new configuration. After updating SRM's configuration and restarting SRM, the old listener can be turned off. Non-inter-broker listeners can be configured with the dynamic configuration API of Kafka, this way not every listener change has to be followed by a restart.

CDPD-11709: Blacklisted topics appear in the list of replicated topics

If a topic was originally replicated but was later excluded for replication, 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, this topic will be included in the response. Additionally, the excluded topic will also be visible in the SMM UI. However, its Partitions and Consumer Groups will be 0, its Throughput, Replication Latency and Checkpoint Latency will show N/A.

None

CDPD-18300: SRM resolves configuration provider references in its internal configuration topic

SRM saves its internal configuration topic with fully resolved properties. This means that even configuration provider references are resolved. Sensitive information can be emitted into the configuration topic this way.

None

CDPD-22094: The SRM service role displays as healthy, but no metrics are processed

The SRM service role might encounter errors that make metrics processing impossible. An example of this is when the target Kafka cluster is not reachable. The SRM service role does not automatically stop or recover if such an error is encountered. It continues to run and displays as healthy in Cloudera Manager. Metrics, however, are not processed. In addition, no new data is displayed in SMM for the replications.

1. Ensure that all clusters are available and are in a healthy state.
2. Restart SRM.

CDPD-22389: The SRM driver role displays as healthy, but replication fails

During startup, the SRM driver role might encounter errors that make data replication impossible. An example of this is when one of the clusters added for replication is not reachable. The SRM driver role does not automatically stop or recover if such an error is encountered. It will start up, continue to run, and display as healthy in Cloudera Manager. Replication, however, will not happen.

1. Ensure that all clusters are available and are in a healthy state.
2. Restart SRM.

CDPD-23683: The replication status reported by the SRM service role for healthy replications is flaky

The replication status reported by the SRM service role is flaky. The replication status might change between active and inactive frequently even if the replication is healthy. This status is also reflected in SMM on the replications tab.

None

Limitations**SRM cannot replicate Ranger authorization policies to or from Kafka clusters**

Due to a limitation in the Kafka-Ranger plugin, SRM cannot replicate Ranger policies to or from clusters that are configured to use Ranger for authorization. If you are using SRM to replicate data to or from a cluster that uses Ranger, disable authorization policy synchronization in SRM. This can be achieved by clearing the Sync Topic Acls Enabled (sync.topic.acls.enabled) checkbox.

Known Issues 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 an intermittent and 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.

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.

Known issues in Streams Messaging Manager

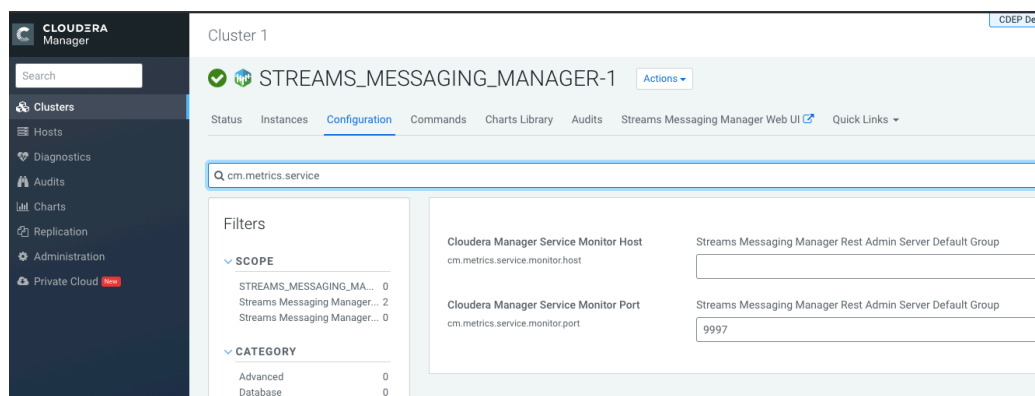
Learn about the known issues in Schema Registry, the impact or changes to the functionality, and the workaround.

CDPD-17361: Consumer Group metrics does not appear in SMM

SMM emits consumer metrics to Cloudera Manager and reads the consumer metrics back from Cloudera Manager. For SMM to work correctly, you must configure the Cloudera Manager Service Monitor Host property. This property is also known as `cm.metrics.service.monitor.host`.

To resolve:

1. Go to the Configuration tab for SMM.
2. Search for `cm.metrics.service.monitor.host`.



3. In the Cloudera Manager Service Monitor Host field, specify the host name of the Service Monitor host.
4. Save the update.

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). 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-separated list of brokers>
```

OPSAPS-59828: SMM cannot connect to Schema Registry when TLS is enabled

When TLS is enabled, SMM by default cannot properly connect to Schema Registry.

As a result, when viewing topics in the SMM Data Explorer with the deserializer key or value set to Avro, the following error messages are shown:

- Error deserializing key/value for partition [***PARTITION***] at offset [***OFFSET***]. If needed, please seek past the record to continue consumption.
- Failed to fetch value schema versions for topic : '[***TOPIC***]'.

In addition, the following certificate error will also be present in the SMM log:

- javax.net.ssl.SSLHandshakeException: PKIX path building failed:...

Workaround: Additional security properties must be set for SMM.

1. In Cloudera Manager, select the SMM service.
2. Go to Configuration.
3. Find and configure the SMM_JMX_OPTS property.

Add the following JVM SSL properties:

- Djavax.net.ssl.trustStore=[***SMM TRUSTSTORE LOCATION***]
- Djavax.net.ssl.trustStorePassword=[***PASSWORD***]

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.

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 only the existing jobs. New jobs started after the move are not affected.

For any existing jobs that have the incorrect JobHistory Server URL, allow the jobs to roll off the history over time. For new jobs, make sure that all clients have the updated mapred-site.xml file 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.address and does not support using the wildcard 0.0.0.0 address.

Set the address, in the form host:port, either in the client-side configuration, or on the command line when you submit the job.

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](#)

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

Yarn Application Log Aggregation will fail for any YARN job (MapReduce, Tez, Spark, and so on) 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 is 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 Resource Managers

When YARN High Availability is enabled there are multiple Resource Managers. 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 the 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.

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:cdwra` which is less secure.

COMPX-5240: Restarting parent queue does not restart child queues in weight mode

When a dynamic auto child creation enabled parent queue is stopped in weight mode, its static and dynamically created child queues are also stopped. However, when the dynamic auto child creation enabled parent queue is restarted, its child queues remain stopped. In addition, the dynamically created child queues cannot be restarted manually through the YARN Queue Manager UI either.

Delete the dynamic auto child creation enabled parent queue. This action also deletes all its child queues, both static and dynamically created child queues, including the stopped dynamic queues. Then recreate the parent queue, enable the dynamic auto child creation feature for it and add any required static child queues.

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 validation failed:java.io.IOException:
Failed to re-init queues : Parent queue 'root.test2' have children queue used mixed of weight
mode, percentage and absolute mode, it is not allowed, please do a double check, details:
{Queue=root.test2.test2childNew, label= uses weight mode}. {Queue=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 implementation, if there is an existing managed queue in Relative mode, then conversion to Weight mode is not 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).

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.

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:cdw which is less secure.

COMPX-8687: Missing access check for getAppAttempts

When the Job ACL feature is enabled using Cloudera Manager (YARN Configuration Enable 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 v2 (ATSV2)
- Container Resizing
- Distributed or Centralized Allocation of Opportunistic Containers
- Distributed Scheduling
- Docker on YARN (DockerContainerExecutor) on Data Hub clusters
- Dynamic Resource Pools
- Fair Scheduler
- GPU support for Docker
- Hadoop Pipes
- Native Services
- Pluggable Scheduler Configuration
- Queue Priority Support
- Reservation REST APIs
- Resource Estimator Service
- Resource Profiles
- (non-Zookeeper) ResourceManager State Store
- Rolling Log Aggregation
- Shared Cache
- YARN Federation

- Moving jobs between queues

Technical Service Bulletins

TSB 2021-539: Capacity Scheduler queue pending metrics can become negative in certain production workload scenarios causing blocked queues

The pending metrics of Capacity Scheduler queues can become negative in certain production workload scenarios.

Once this metric becomes negative, the scheduler is unable to schedule any further resource requests on the specific queue. As a result, new applications are stuck in the ACCEPTED state unless YARN ResourceManager is restarted or failed-over.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-539: Capacity Scheduler queue pending metrics can become negative in certain production workload scenarios causing blocked queues](#)

Known Issues in Apache Zeppelin

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

BUG-125263: Zeppelin service move fails on clusters upgraded from HDP3.1.5

Resolve the circular symlink issue on the Zeppelin node by linking the conf directory to a new directory under /etc/zeppelin:

- `# mkdir -p /etc/zeppelin/<version>/0`
- `# rm /usr/hdp/<version>/zeppelin/conf`
- `# ln -s /etc/zeppelin/<version>/0 /usr/hdp/<version>/zeppelin/conf`

Where version is the HDP version. For example, 7.1.x-yyy. Restart the Zeppelin server in Ambari.

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_group}}]
```

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 7.1.4

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 7.1.4.

Behavioral Changes in Apache Kudu

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

Summary:

Changing the default value of RPC queue length.

Details:

Previous behavior:

The default value of RPC queue length was 50 for Kudu masters.

New behavior:

The default value of RPC queue length is 100 for Kudu masters.

Behavioral Changes in Apache Phoenix

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

Summary:

As a consequence of including the Apache Phoenix fix PHOENIX-5213, some libraries that are in the old shaded phoenix-client JARs are included in the new JAR without relocation. These libraries are shaded in the new JAR. For example, com.google.protobuf package is shaded in the new JAR.

New behavior and action required:

If you use any of these JARs in your current deployment, you must explicitly add the newly shaded dependencies to your applications.

Behavioral Changes in Ranger

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

Cloudera ID

CDPD-4141

Apache component

Ranger

Apache JIRA:

N/A

Summary:

Default location for ranger audit data collected by Solr changed to local file system from HDFS

Details:**Previous behavior:**

The default value of the index storage in HDFS is /solr-infra.

New behavior:

The default value of the index storage in the local file system is /var/lib/solr-infra.

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.

Cloudera ID

OPSAPS-57361

Apache component

Solr

Apache JIRA:

N/A

Summary:

Upgrade behavior changed

Details:**Previous behavior:**

The solr-upgrade.sh script used to perform a "Solr bootstrap collections" step.

New behavior:

The solr-upgrade.sh script omits the collections bootstrap step, you need to bootstrap the recreated empty collections after the upgrade, but before reindexing collections.

Deprecation Notices In Cloudera Runtime 7.1.4

Certain features and functionalities have been removed or deprecated in Cloudera Runtime 7.1.4. 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 Apache Kudu are deprecated or removed in Cloudera Runtime 7.1.4. You must review these changes along with the information about the features in Apache 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 Kafka are deprecated or removed in Cloudera Runtime 7.1.4. 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 kafka-configs and kafka-reassign-partitions. Cloudera recommends that you use the --bootstrap-server option instead.

Deprecation Notices in Apache HBase

Certain features and functionality in Hbase are deprecated or removed in Cloudera Runtime 7.1.4. You must review these changes along with the information about the features in HBase that will be removed or deprecated in a future release.

Use this list to understand some of the deprecated items and incompatibilities if you are upgrading from HDP 2.x or CDH 5.x to CDP.

Known Incompatibilities when Upgrading from CDH and HDP

Cloudera Runtime uses Apache HBase 2.x.x whereas CDH 5.x and HDP 2.x uses Apache HBase 1.x.



Important: Some APIs that are listed as deprecated, but these APIs do not block your upgrade. You must stop using the deprecated APIs in your existing applications after upgrade, and not use these APIs in new development.

List of Major Changes

- HBASE-16189 and HBASE-18945: You cannot open the Cloudera Runtime HFiles in CDH or HDP.
- HBASE-18240: Changed the ReplicationEndpoint Interface.
- The Dynamic Jars Directory property `hbase.dynamic.jars.dir` is disabled by default. If you want to enable dynamic classloading, you can use the `hbase.dynamic.jars.dir` property in Cloudera Manager to change the default `${hbase.rootdir}/lib` directory to some other location, preferably a location on HDFS. This property is flagged by Cloudera Manager as deprecated when you upgrade to CDP because the property is incompatible with HBase on cloud deployments. If you are using HBase with HDFS storage, you can ignore this warning, and keep using the `hbase.use.dynamic.jars` feature.

Co-processor API changes

- HBASE-16769: Deprecated Protocol Buffers references from `MasterObserver` and `RegionServerObserver`.
- HBASE-17312: [JDK8] Use default method for `Observer` Coprocessors. The interface classes of `BaseMasterAndRegionObserver`, `BaseMasterObserver`, `BaseRegionObserver`, `BaseRegionServerObserver` and `BaseWALObserver` uses JDK8's 'default' keyword to provide empty and no-op implementations.
- Interface `HTableInterface` introduces following changes to the methods listed below:

[#] interface `CoprocessorEnvironment`

Change	Result
Abstract method <code>getTable (TableName)</code> has been removed.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>getTable (TableName, ExecutorService)</code> has been removed.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.

- Public Audience

The following tables describes the coprocessor changes:

[#] class `CoprocessorRpcChannel` (1)

Change	Result
This class has become interface.	A client program may be interrupted by <code>IncompatibleClassChangeError</code> or <code>InstantiationException</code> exception depending on the usage of this class.

Class `CoprocessorHost<E>`

Classes that were Audience Private but were removed:

Change	Result
Type of field <code>coprocessors</code> has been changed from <code>java.util.SortedSet<E></code> to <code>org.apache.hadoop.hbase.util.SortedList<E></code> .	A client program may be interrupted by <code>NoSuchFieldError</code> exception.

MasterObserver changes

The following changes are introduced to the `MasterObserver` interface:

[#] interface `MasterObserver` (14)

Change	Result
Abstract method <code>voidpostCloneSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostCreateTable (ObserverContext<MasterCoproprocessorEnvironment>, HTableDescriptor, HRegionInfo[])</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostDeleteSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostGetTableDescriptors (ObserverContext<MasterCoproprocessorEnvironment>, List<HTableDescriptor>)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostModifyTable (ObserverContext<MasterCoproprocessorEnvironment>, TableName, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostRestoreSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpreCloneSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpreCreateTable (ObserverContext<MasterCoproprocessorEnvironment>, HTableDescriptor, HRegionInfo[])</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpreDeleteSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpreGetTableDescriptors (ObserverContext<MasterCoproprocessorEnvironment>, List<TableName>, List<HTableDescriptor>)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpreModifyTable (ObserverContext<MasterCoproprocessorEnvironment>, TableName, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpreRestoreSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpreSnapshot (ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.

RegionObserver interface changes

The following changes are introduced to the `RegionObserver` interface.

[#] interface `RegionObserver` (13)

Change	Result
Abstract method <code>voidpostCloseRegionOperation (ObserverContext<RegionCoproprocessorEnvironment>, HRegion.Operation)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostCompactSelection (ObserverContext<RegionCoproprocessorEnvironment>, Store, ImmutableList<StoreFile>)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.

Abstract method voidpostCompactSelection (ObserverContext<RegionCoproprocessorEnvironment>, Store, ImmutableList<StoreFile>, CompactionRequest) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method voidpostGetClosestRowBefore (ObserverContext<RegionCoproprocessorEnvironment>, byte[], byte[], Result) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method DeleteTrackerpostInstantiateDeleteTracker (ObserverContext<RegionCoproprocessorEnvironment>, DeleteTracker) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method voidpostSplit (ObserverContext<RegionCoproprocessorEnvironment>, HRegion, HRegion) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method voidpostStartRegionOperation (ObserverContext<RegionCoproprocessorEnvironment>, HRegion.Operation) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method StoreFile.ReaderpostStoreFileReaderOpen (ObserverContext<RegionCoproprocessorEnvironment>, FileSystem, Path, FSDataInputStreamWrapper, long, CacheConfig, Reference, StoreFile.Reader) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method voidpostWALRestore (ObserverContext<RegionCoproprocessorEnvironment>, HRegionInfo, HLogKey, WALEdit) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method InternalScannerpreFlushScannerOpen (ObserverContext<RegionCoproprocessorEnvironment>, Store, KeyValueScanner, InternalScanner) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method voidpreGetClosestRowBefore (ObserverContext<RegionCoproprocessorEnvironment>, byte[], byte[], Result) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method StoreFile.ReaderpreStoreFileReaderOpen (ObserverContext<RegionCoproprocessorEnvironment>, FileSystem, Path, FSDataInputStreamWrapper, long, CacheConfig, Reference, StoreFile.Reader) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method voidpreWALRestore (ObserverContext<RegionCoproprocessorEnvironment>, HRegionInfo, HLogKey, WALEdit) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

WALObserver interface changes

The following changes are introduced to the WALObserver interface:

[#] interface WALObserver

Change	Result
Abstract method voidpostWALWrite (ObserverContext<WALCoproprocessorEnvironment>, HRegionInfo, HLogKey, WALEdit) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method booleanpreWALWrite (ObserverContext<WALCoproprocessorEnvironment>, HRegionInfo, HLogKey, WALEdit) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

Scheduler changes

Following methods are now changed to abstract:

[#]class RpcScheduler (1)

Change	Result
--------	--------

Abstract method void dispatch (CallRunner) has been removed from this class.	A client program may be interrupted by NoSuchMethodError exception.
--	---

[#] RpcScheduler.dispatch (CallRunner p1) [abstract] : void 1

org/apache/hadoop/hbase/ipc/RpcScheduler.dispatch:(Lorg/apache/hadoop/hbase/ipc/CallRunner;)V

Change	Result
Return value type has been changed from void to boolean.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

The following abstract methods have been removed:

[#]interface PriorityFunction (2)

Change	Result
Abstract method longgetDeadline (RPCProtos.RequestHeader, Message) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method int getPriority (RPCProtos.RequestHeader, Message) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

Server API changes

[#] class RpcServer (12)

Change	Result
Type of field CurCall has been changed from java.lang.ThreadLocal<RpcServer.Call> to java.lang.ThreadLocal<RpcCall>.	A client program may be interrupted by NoSuchFieldError exception.
Abstract method int getNumOpenConnections () has been added to this class.	This class became abstract and a client program may be interrupted by InstantiationException exception.
Field callQueueSize of type org.apache.hadoop.hbase.util.Counter has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field connectionList of type java.util.List<RpcServer.Connection> has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field maxIdleTime of type int has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field numConnections of type int has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field port of type int has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field purgeTimeout of type long has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field responder of type RpcServer.Responder has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field socketSendBufferSize of type int has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field thresholdIdleConnections of type int has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.

Following abstract methods are removed:

Change	Result
Abstract method Pair<Message,CellScanner>call (BlockingService, Descriptors.MethodDescriptor, Message, CellScanner, long, MonitoredRPCHandler) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

Replication and WAL changes

HBASE-18733: WALKey has been purged completely. Following are the changes to the WALKey:

[#] classWALKey (8)

Change	Result
Access level of field clusterIds has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field compressionContext has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field encodedRegionName has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field tablename has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field writeTime has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.

Following fields have been removed:

Change	Result
Field LOG of type <code>org.apache.commons.logging.Log</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field VERSION of type <code>WALKey.Version</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>logSeqNum</code> of type <code>long</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.

Admin Interface API changes

You cannot administer a CDP Runtime Data Hub cluster using a client that includes `RelicationAdmin`, `ACC`, Thrift and REST usage of Admin ops. Methods returning protobufs have been changed to return POJOs instead. Returns have changed from void to Future for async methods. HBASE-18106 - `Admin.listProcedures` and `Admin.listLocks` were renamed to `getProcedures` and `getLocks`. MapReduce makes use of Admin doing following `admin.getClusterStatus()` to calculate Splits.

- Thrift usage of Admin API:

```
compact(ByteBuffer) createTable(ByteBuffer, List<ColumnDescriptor>) deleteTable(ByteBuffer) disableTable(ByteBuffer) enableTable(ByteBuffer) getTableNames() majorCompact(ByteBuffer)
```

- REST usage of Admin API:

```
hbase-rest org.apache.hadoop.hbase.rest RootResource getTableList() TableName[] tableNames = servlet.getAdmin().listTableNames(); SchemaResource delete(UriInfo) Admin admin = servlet.getAdmin(); update(TableSchemaModel, boolean, UriInfo) Admin admin = servlet.getAdmin(); StorageClusterStatusResource get(UriInfo) ClusterStatus status = servlet.getAdmin().getClusterStatus(); StorageClusterVersionResource get(UriInfo) model.setVersion(servlet.getAdmin().getClusterStatus().getHBaseVersion()); TableResource exists() return servlet.getAdmin().tableExists(TableName.valueOf(table));
```

[#] interface Admin (9)

Following are the changes to the Admin interface:

Change	Result
--------	--------

Abstract method createTableAsync (HTableDescriptor, byte[][]) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method disableTableAsync (TableName) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method enableTableAsync (TableName) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method getCompactionState (TableName) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method getCompactionStateForRegion (byte[]) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method isSnapshotFinished (HBaseProtos.SnapshotDescription) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method snapshot (String, TableName, HBaseProtos.SnapshotDescription.Type) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method snapshot (HBaseProtos.SnapshotDescription) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method takeSnapshotAsync (HBaseProtos.SnapshotDescription) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

[#] Admin.createTableAsync (HTableDescriptor p1, byte[][] p2) [abstract] : void 1

org/apache/hadoop/hbase/client/Admin.createTableAsync:(Lorg/apache/hadoop/hbase/HTableDescriptor;[[B)V

Change	Result
Return value type has been changed from void to java.util.concurrent.Future<java.lang.Void>.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

[#] Admin.disableTableAsync (TableName p1) [abstract] : void 1

org/apache/hadoop/hbase/client/Admin.disableTableAsync:(Lorg/apache/hadoop/hbase/TableName;)V

Change	Result
Return value type has been changed from void to java.util.concurrent.Future<java.lang.Void>.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

Admin.enableTableAsync (TableName p1) [abstract] : void 1

org/apache/hadoop/hbase/client/Admin.enableTableAsync:(Lorg/apache/hadoop/hbase/TableName;)V

Change	Result
Return value type has been changed from void to java.util.concurrent.Future<java.lang.Void>.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

Admin.enableTableAsync (TableName p1) [abstract] : void 1

org/apache/hadoop/hbase/client/Admin.getCompactionState:(Lorg/apache/hadoop/hbase/TableName;)Lorg/apache/hadoop/hbase/protobuf/generated/AdminProtos\$GetRegionInfoResponse\$CompactionState;

Change	Result
Return value type has been changed from org.apache.hadoop.hbase.protobuf.generated.AdminProtos.GetRegionInfoResponse.CompactionState to CompactionState.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

[#] Admin.getCompactionStateForRegion (byte[] p1) [abstract] : AdminProtos.GetRegionInfoResponse.CompactionState 1

org/apache/hadoop/hbase/client/Admin.getCompactionStateForRegion:([B)Lorg/apache/hadoop/hbase/protobuf/generated/AdminProtos\$GetRegionInfoResponse\$CompactionState;

Change	Result
Return value type has been changed from org.apache.hadoop.hbase.protobuf.generated.AdminProtos.GetRegionInfoResponse.CompactionState to CompactionState.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

HTableDescriptor and HColumnDescriptor changes

HTableDescriptor and HColumnDescriptor has become interfaces and you can create it through Builders. HCD has become CFD. It no longer implements writable interface. package org.apache.hadoop.hbase.

[#] class HColumnDescriptor (1)

Change	Result
Removed super-interface org.apache.hadoop.io.WritableComparable<HColumnDescriptor>.	A client program may be interrupted by NoSuchMethodError exception.

class HTableDescriptor (3)

Change	Result
Removed super-interface org.apache.hadoop.io.WritableComparable<HTableDescriptor>.	A client program may be interrupted by NoSuchMethodError exception.
Field META_TABLEDESC of type HTableDescriptor has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.

[#] HTableDescriptor.getColumnFamilies () : HColumnDescriptor[] (1)

org/apache/hadoop/hbase/HTableDescriptor.getColumnFamilies:()[Lorg/apache/hadoop/hbase/HColumnDescriptor;

[#] class HColumnDescriptor (1)

Change	Result
Return value type has been changed from HColumnDescriptor[] to client.ColumnFamilyDescriptor[].	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

[#] interface Table (4)

Change	Result
Abstract method batch (List<?>) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method batchCallback (List<?>, Batch.Callback<R>) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method getWriteBufferSize () has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method setWriteBufferSize (long) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

Deprecated buffer methods

- LockTimeoutException and OperationConflictException classes have been removed.

class OperationConflictException (1)

Result	Result
This class has been removed.	A client program may be interrupted by NoClassDefFoundError exception.

class class LockTimeoutException (1)

Change Result This class has been removed. A client program may be interrupted by NoClassDefFoundError exception.

Filter API changes

Following methods have been removed: package org.apache.hadoop.hbase.filter

[#] class Filter (2)

Result	Result
Abstract method getNextKeyHint (KeyValue) has been removed from this class.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method transform (KeyValue) has been removed from this class.	A client program may be interrupted by NoSuchMethodError exception.

- HBASE-12296: Filters should work with ByteBufferedCell.
- HConnection is removed in Cloudera Runtime.
- RegionLoad and ServerLoad internally moved to shaded Protocol Buffers.

[#] class RegionLoad (1)

Result	Result
Type of field regionLoadPB has been changed from protobuf.generated.ClusterStatusProtos.RegionLoad to shaded.protobuf.generated.ClusterStatusProtos.RegionLoad.	A client program may be interrupted by NoSuchFieldError exception.

[#] interface AccessControlConstants (3)

Result	Result
Field OP_ATTRIBUTE_ACL_STRATEGY of type java.lang.String has been removed from this interface.	A client program may be interrupted by NoSuchFieldError exception.
Field OP_ATTRIBUTE_ACL_STRATEGY_CELL_FIRST of type byte[] has been removed from this interface.	A client program may be interrupted by NoSuchFieldError exception.
Field OP_ATTRIBUTE_ACL_STRATEGY_DEFAULT of type byte[] has been removed from this interface.	A client program may be interrupted by NoSuchFieldError exception.

[#] ServerLoad.getNumberOfRequests () : int 1

org/apache/hadoop/hbase/ServerLoad.getNumberOfRequests():I

Result	Result
Return value type has been changed from int to long.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

[#] ServerLoad.getNumberOfRequests () : int 1

org/apache/hadoop/hbase/ServerLoad.getReadRequestsCount():I

Result	Result
Return value type has been changed from int to long.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

[#] ServerLoad.getTotalNumberOfRequests () : int 1

org/apache/hadoop/hbase/ServerLoad.getTotalNumberOfRequests():I

Result	Result
Return value type has been changed from int to long.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

[#]ServerLoad.getWriteRequestsCount () : int 1

org/apache/hadoop/hbase/ServerLoad.getWriteRequestsCount:()I

Result	Result
Return value type has been changed from int to long.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.

[#]class HConstants (6)

Result	Result
Field DEFAULT_HBASE_CONFIG_READ_ZOOKEEPER_CONFIG of type boolean has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field HBASE_CONFIG_READ_ZOOKEEPER_CONFIG of type java.lang.String has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field REPLICATION_ENABLE_DEFAULT of type boolean has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field REPLICATION_ENABLE_KEY of type java.lang.String has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field ZOOKEEPER_CONFIG_NAME of type java.lang.String has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.
Field ZOOKEEPER_USEMULTI of type java.lang.String has been removed from this class.	A client program may be interrupted by NoSuchFieldError exception.

HBASE-18732: [compat 1-2] HBASE-14047 removed Cell methods without deprecation cycle.

[#]interface Cell 5

Result	Result
Abstract method getFamily () has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method getMvccVersion () has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method getQualifier () has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method getRow () has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.
Abstract method getValue () has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

HBASE-18795:Expose KeyValue.getBuffer() for tests alone. Allows KV#getBuffer in tests only that was deprecated previously.

Region scanner changes

[#]interface RegionScanner (1)

Result	Result
Abstract method boolean nextRow (List<Cell>, int) has been removed from this interface.	A client program may be interrupted by NoSuchMethodError exception.

StoreFile changes

[#] class StoreFile (1)

Result	Result
This class became interface.	A client program may be interrupted by <code>IncompatibleClassChangeError</code> or <code>InstantiationError</code> exception dependent on the usage of this class.

MapReduce changes

HFile*Format has been removed.

ClusterStatus changes

[#] ClusterStatus.getRegionsInTransition () : Map<String,RegionState> 1

org/apache/hadoop/hbase/ClusterStatus.getRegionsInTransition:()Ljava/util/Map;

Result	Result
Return value type has been changed from <code>java.util.Map<java.lang.String,master.RegionState></code> to <code>java.util.List<master.RegionState></code> .	This method has been removed because the return type is part of the method signature. A client program may be interrupted by <code>NoSuchMethodError</code> exception.

Other changes in ClusterStatus include removal of convert methods that were no longer necessary after purge of Protocol Buffers from API.

Purge of Protocol Buffers from API

Protocol Buffers (PB) has been deprecated in APIs.

[#] HBaseSnapshotException.getSnapshotDescription () : HBaseProtos.SnapshotDescription 1

org/apache/hadoop/hbase/snapshot/HBaseSnapshotException.getSnapshotDescription:()Lorg/apache/hadoop/hbase/protobuf/generated/HBaseProtos\$SnapshotDescription;

Result	Result
Return value type has been changed from <code>org.apache.hadoop.hbase.protobuf.generated.HBaseProtos.SnapshotDescription</code> to <code>org.apache.hadoop.hbase.client.SnapshotDescription</code> .	This method has been removed because the return type is part of the method signature. A client program may be interrupted by <code>NoSuchMethodError</code> exception.

HBASE-15609: Remove PB references from Result, DoubleColumnInterpreter and any such public facing class for 2.0. hbase-client-1.0.0.jar, Result.class package org.apache.hadoop.hbase.client

[#] Result.getStats () : ClientProtos.RegionLoadStats 1

org/apache/hadoop/hbase/client/Result.getStats:()Lorg/apache/hadoop/hbase/protobuf/generated/ClientProtos\$RegionLoadStats;

Result	Result
Return value type has been changed from <code>org.apache.hadoop.hbase.protobuf.generated.ClientProtos.RegionLoadStats</code> to <code>RegionLoadStats</code> .	This method has been removed because the return type is part of the method signature. A client program may be interrupted by <code>NoSuchMethodError</code> exception.

PrettyPrinter changes

hbase-server-1.0.0.jar, HFilePrettyPrinter.class package org.apache.hadoop.hbase.io.hfile

Result	Result
--------	--------

Return value type has been changed from void to int.

This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError exception.