

Cloudera Runtime 7.1.1

## Release Notes

Date published: 2020-05-22

Date modified:

# CLOUdera

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

# Legal Notice

© Cloudera Inc. 2024. 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

<b>Overview.....</b>	<b>5</b>
<b>Cloudera Runtime Component Versions.....</b>	<b>5</b>
<b>Using the Cloudera Runtime Maven Repository.....</b>	<b>6</b>
Maven Artifacts for Cloudera Runtime 7.1.1.0.....	6
<b>What's New In Cloudera Runtime 7.1.1.....</b>	<b>24</b>
What's New in Apache Atlas.....	24
What's New in Cruise Control.....	26
What's new in DAS.....	26
What's New in Apache HBase.....	27
What's New in Apache Hadoop HDFS.....	27
What's New in Apache Hive.....	27
What's New in Hue.....	27
What's New in Apache Impala.....	29
What's New in Apache Kafka.....	30
What's New in Apache Knox.....	31
What's New in Apache Kudu.....	31
What's New in Apache Oozie.....	33
What's New in Apache Hadoop Ozone.....	33
What's New in Apache Phoenix.....	34
What's New in Schema Registry.....	34
What's New in Cloudera Search.....	34
What's New in Apache Spark.....	35
What's New in Sqoop.....	35
What's New in Streams Replication Manager.....	35
What's new in Streams Messaging Manager.....	36
What's New in Apache Hadoop YARN.....	36
What's New in Apache ZooKeeper.....	37
<b>Deprecation Notices In Cloudera Runtime 7.1.1.....</b>	<b>37</b>
Deprecation Notices in Apache HBase.....	38
Deprecation notices in Apache Kudu.....	48
Deprecation Notices in Cloudera Search.....	48
Deprecation Notices for Apache Kafka.....	49
<b>Behavioral Changes In Cloudera Runtime 7.1.1.....</b>	<b>49</b>
Behavioral Changes in Cloudera Search.....	50
Behavioral Changes in Apache Hive.....	50
Behavioral Changes in Apache Hadoop YARN.....	51
<b>Fixed Issues In Cloudera Runtime 7.1.1.....</b>	<b>51</b>

Fixed issues in Atlas.....	51
Fixed issues in DAS.....	52
Fixed issues in Hadoop.....	52
Fixed issues in HBase.....	52
Fixed issues in HDFS.....	53
Fixed Issues in Hive.....	53
Fixed Issues in Hue.....	53
Fixed Issues in Kafka.....	53
Fixed Issues in Impala.....	53
Fixed Issues in Kudu.....	54
Fixed issues in Oozie.....	54
Fixed issues in Ozone.....	56
Fixed issues in Phoenix.....	56
Fixed issues in Search.....	56
Fixed issues in Spark.....	57
Fixed issues in Sqoop.....	57
Fixed Issues in Streams Replication Manager.....	57
Fixed Issues in Streams Messaging Manager.....	58
Fixed issues in YARN.....	59
Fixed issues in Zeppelin.....	59
Fixed issues in ZooKeeper.....	59

## **Known Issues In Cloudera Runtime 7.1.1..... 60**


Known Issues in Apache Atlas.....	60
Known issues in Cruise Control.....	63
Known Issues in DAS.....	64
Known Issues in Apache Hadoop.....	65
Known Issues in Apache HBase.....	66
Known Issues in HDFS.....	68
Known Issues in Apache Hive.....	69
Known Issues in Hue.....	72
Known Issues in Apache Impala.....	75
Known Issues in Apache Kafka.....	80
Known Issues in Kerberos.....	84
Known Issues in Apache Knox.....	84
Known Issues in Apache Kudu.....	85
Known Issues in Apache Oozie.....	85
Ozone.....	86
Known Issues in Apache Ranger.....	86
Known Issues in Apache Ranger KMS.....	87
Known Issues in Schema Registry.....	87
Known Issues in Cloudera Search.....	87
Known Issues in Apache Solr.....	93
Known Issues in Apache Spark.....	93
Known Issues in Streams Replication Manager.....	95
Known issues in Streams Messaging Manager.....	97
Known Issues for Apache Sqoop.....	98
Known Issues in MapReduce and YARN.....	99
Known Issues in Apache Zeppelin.....	102
Known Issues in Apache ZooKeeper.....	102

## Overview

Welcome to the Cloudera Runtime Release Notes. This document provides you with the latest information about Cloudera Runtime 7.1.1. This document includes improvements and describes new features, bug fixes, tech previews and more. For detailed information about the runtime components themselves, see [Cloudera documentation](#).

## Cloudera Runtime Component Versions

List of the official Apache component versions for Cloudera Runtime. To know the Apache component versions for compatibility with other applications, you must be familiar with the latest Apache component versions in Cloudera Runtime. You should also be aware of the available Technical Preview components and use them only in a testing environment. Apache versions of Cloudera Runtime 7.1.1. components.

Component	Version
Apache Atlas	2.0.0
Apache Avro	1.8.2
Apache Hadoop	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 Oozie	5.1.0
Apache ORC	1.5.1
Apache Ozone	0.5.0
Apache Parquet	1.10.99
Apache Phoenix	5.0.0
Apache Ranger	2.0.0
Schema Registry	0.8.1
Apache Solr	8.4.1
Apache Spark	 <b>Note:</b> 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 7.1.1 is based on Apache Spark 2.4.5, not 2.4.0.
	2.4.5
Apache Sqoop	1.4.7
Apache Tez	0.9.1
Apache ZooKeeper	3.5.5
Apache Zeppelin	0.8.2
Cruise Control	2.0.100
DAS	1.4.2

Component	Version
HBase Indexer	1.5.0
HDFS	3.1.1
Hue	4.5.0
Streams Messaging Manager	2.1.0
Streams Replication Manager	1.0.0
YARN	3.1.1

## Using the Cloudera Runtime Maven Repository

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](https://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 7.1.1.0

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

Project	groupId	artifactId	version
Apache Atlas	org.apache.atlas	atlas-authorization	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-classification-updater	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-client-common	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-client-v1	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-client-v2	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-common	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-distro	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-docs	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-graphdb-api	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-graphdb-common	2.0.0.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.atlas	atlas-graphdb-janus	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-intg	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-janusgraph-hbase2	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-notification	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-plugin-classloader	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-repository	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-server-api	2.0.0.7.1.1.0-565
	org.apache.atlas	atlas-testtools	2.0.0.7.1.1.0-565
	org.apache.atlas	hbase-bridge	2.0.0.7.1.1.0-565
	org.apache.atlas	hbase-bridge-shim	2.0.0.7.1.1.0-565
	org.apache.atlas	hbase-testing-util	2.0.0.7.1.1.0-565
	org.apache.atlas	hdfs-model	2.0.0.7.1.1.0-565
	org.apache.atlas	hive-bridge	2.0.0.7.1.1.0-565
	org.apache.atlas	hive-bridge-shim	2.0.0.7.1.1.0-565
	org.apache.atlas	impala-bridge	2.0.0.7.1.1.0-565
	org.apache.atlas	impala-bridge-shim	2.0.0.7.1.1.0-565
	org.apache.atlas	impala-hook-api	2.0.0.7.1.1.0-565
	org.apache.atlas	kafka-bridge	2.0.0.7.1.1.0-565
	org.apache.atlas	navigator-to-atlas	2.0.0.7.1.1.0-565
	org.apache.atlas	sqoop-bridge	2.0.0.7.1.1.0-565
	org.apache.atlas	sqoop-bridge-shim	2.0.0.7.1.1.0-565
Apache Avro	org.apache.avro	avro	1.8.2.7.1.1.0-565
	org.apache.avro	avro-compiler	1.8.2.7.1.1.0-565
	org.apache.avro	avro-ipc	1.8.2.7.1.1.0-565
	org.apache.avro	avro-mapred	1.8.2.7.1.1.0-565
	org.apache.avro	avro-maven-plugin	1.8.2.7.1.1.0-565
	org.apache.avro	avro-protobuf	1.8.2.7.1.1.0-565
	org.apache.avro	avro-service-archetype	1.8.2.7.1.1.0-565
	org.apache.avro	avro-thrift	1.8.2.7.1.1.0-565
	org.apache.avro	avro-tools	1.8.2.7.1.1.0-565
	org.apache.avro	trevni-avro	1.8.2.7.1.1.0-565
	org.apache.avro	trevni-core	1.8.2.7.1.1.0-565
Apache Calcite	org.apache.calcite	calcite-babel	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-cassandra	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-core	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-druid	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-elasticsearch	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-example-csv	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-example-function	1.19.0.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.calcite	calcite-file	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-geode	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-linq4j	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-mongodb	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-pig	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-piglet	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-plus	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-server	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-spark	1.19.0.7.1.1.0-565
	org.apache.calcite	calcite-splunk	1.19.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica	1.10.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica-core	1.10.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica-metrics	1.10.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica-metrics-dropwizardmetrics3	1.10.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica-noop-driver	1.10.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica-server	1.10.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica-standalone-server	1.10.0.7.1.1.0-565
	org.apache.calcite.avatica	avatica-tck	1.10.0.7.1.1.0-565
Apache Crunch	org.apache.crunch	crunch-archetype	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-contrib	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-core	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-examples	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-hbase	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-hive	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-scrunch	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-spark	0.11.0.7.1.1.0-565
	org.apache.crunch	crunch-test	0.11.0.7.1.1.0-565
Apache Druid	org.apache.druid	druid-aws-common	0.15.1.7.1.1.0-565
	org.apache.druid	druid-benchmarks	0.15.1.7.1.1.0-565
	org.apache.druid	druid-console	0.15.1.7.1.1.0-565
	org.apache.druid	druid-core	0.15.1.7.1.1.0-565
	org.apache.druid	druid-gcp-common	0.15.1.7.1.1.0-565
	org.apache.druid	druid-hll	0.15.1.7.1.1.0-565
	org.apache.druid	druid-indexing-hadoop	0.15.1.7.1.1.0-565
	org.apache.druid	druid-indexing-service	0.15.1.7.1.1.0-565
	org.apache.druid	druid-integration-tests	0.15.1.7.1.1.0-565
	org.apache.druid	druid-processing	0.15.1.7.1.1.0-565
	org.apache.druid	druid-server	0.15.1.7.1.1.0-565
	org.apache.druid	druid-services	0.15.1.7.1.1.0-565



Project	groupId	artifactId	version
	org.apache.druid	druid-sql	0.15.1.7.1.1.0-565
	org.apache.druid	extendedset	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hadoop-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hbase-security	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hbase-bloom-filter	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hbase-datasketches	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hbase2-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hive-examples	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hive-google-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hdfs-storage	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-histogram	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-kafka-eight	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-kafka-extraction-namespace	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-kafka-indexing-service	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-kerberos	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-kinesis-indexing-service	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-lookups-cached-global	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-lookups-cached-single	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-morc-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-parquet-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-protobuf-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-rs3-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-stats	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-sql-metadata-storage	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-sql-metadata-storage	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-client-sslcontext	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-metrics-emitter	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-metrics-emitter	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-storage	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-files-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-bcount	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-light-simple-consumer	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-sketch	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-average-query	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk-emitter	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-splunk	0.15.1.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.druid.extensions	druid-hacktribq	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-extensions	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-min-max	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-columns	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-contrib	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-contrib	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-view-maintenance	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-view-selection	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-data-storage	0.15.1.7.1.1.0-565
	org.apache.druid.extensions	druid-hacktribq-contrib	0.15.1.7.1.1.0-565
GCS Connector	com.google.cloud.bigtable	gcs-connector	1.9.10.7.1.1.0-565
	com.google.cloud.bigtable	gcs-connector	1.9.10.7.1.1.0-565
	com.google.cloud.bigtable	gcs-connector	1.9.10.7.1.1.0-565
	com.google.cloud.bigtable	gcs-connector	1.9.10.7.1.1.0-565
	com.google.cloud.bigtable	gcs-connector	1.9.10.7.1.1.0-565
Apache Hadoop	org.apache.hadoop	hadoop-aliyun	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-annotations	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-archive-logs	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-archives	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-assemblies	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-auth	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-aws	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-azure	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-azure-datalake	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-build-tools	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-client	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-client-api	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-client-integration-tests	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-client-minicluster	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-client-runtime	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-cloud-storage	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-common	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-datajoin	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-distcp	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-extras	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-fs2img	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-gridmix	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-client	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-common	0.5.0.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-hdds-config	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-container-service	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-docs	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-hadoop-dependency-client	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-hadoop-dependency-server	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-hadoop-dependency-test	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-server-framework	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-server-scm	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdds-tools	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-hdfs	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-hdfs-client	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-hdfs-httpfs	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-hdfs-native-client	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-hdfs-nfs	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-hdfs-rbf	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-kafka	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-kms	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-app	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-common	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-core	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-hs	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-hs-plugins	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-jobclient	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-nativetask	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-shuffle	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-client-uploader	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-mapreduce-examples	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-maven-plugins	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-minicluster	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-minikdc	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-nfs	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-openstack	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-client	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-common	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-csi	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-datanode	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-filesystem	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-filesystem-lib-current	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-filesystem-lib-legacy	0.5.0.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-ozone-insight	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-integration-test	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-network-tests	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-ozone-manager	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-recon	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-reconcodegen	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-s3gateway	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-tools	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-ozone-upgrade	0.5.0.7.1.1.0-565
	org.apache.hadoop	hadoop-resourceestimator	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-rumen	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-sls	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-streaming	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-tools-dist	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-api	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-applications-distributedshell	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-applications-unmanaged-am-launcher	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-client	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-common	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-registry	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-applicationhistoryservice	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-common	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-nodemanager	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-resourcemanager	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-router	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-sharedcachemanager	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-tests	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-timeline-pluginstorage	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-timelineservice	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-client	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-common	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-server-2	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-tests	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-server-web-proxy	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-services-api	3.1.1.7.1.1.0-565
	org.apache.hadoop	hadoop-yarn-services-core	3.1.1.7.1.1.0-565
	org.apache.hadoop	mini-chaos-tests	0.5.0.7.1.1.0-565
Apache HBase	org.apache.hbase	hbase-annotations	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-checkstyle	2.2.3.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.hbase	hbase-client	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-client-project	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-common	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-endpoint	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-examples	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-external-blockcache	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-hadoop-compat	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-hadoop2-compat	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-hbtop	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-http	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-it	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-mapreduce	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-metrics	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-metrics-api	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-procedure	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-protocol	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-protocol-shaded	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-replication	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-resource-bundle	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-rest	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-rsgroup	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-server	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-shaded-client	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-shaded-client-byo-hadoop	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-shaded-client-project	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-shaded-mapreduce	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-shaded-testing-util	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-shaded-testing-util-tester	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-shell	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-testing-util	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-thrift	2.2.3.7.1.1.0-565
	org.apache.hbase	hbase-zookeeper	2.2.3.7.1.1.0-565
	org.apache.hbase.connector.kafka	hbase-kafka-model	1.0.0.7.1.1.0-565
	org.apache.hbase.connector.kafka	hbase-kafka-proxy	1.0.0.7.1.1.0-565
	org.apache.hbase.connector.spark	hbase-spark	1.0.0.7.1.1.0-565
	org.apache.hbase.connector.spark	hbase-spark-it	1.0.0.7.1.1.0-565
	org.apache.hbase.connector.spark	hbase-spark-protocol	1.0.0.7.1.1.0-565
	org.apache.hbase.connector.spark	hbase-spark-protocol-shaded	1.0.0.7.1.1.0-565
	org.apache.hbase.filesystem	hbase-filesystem	1.0.0.7.1.1.0-565

Project	groupId	artifactId	version
Apache Hive	org.apache.hive	hive-accumulo-handler	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-beeline	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-classification	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-cli	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-common	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-contrib	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-druid-handler	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-exec	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-hbase-handler	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-hplsql	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-jdbc	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-jdbc-handler	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-kryo-registrator	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-kudu-handler	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-llap-client	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-llap-common	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-llap-ext-client	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-llap-server	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-llap-tez	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-metastore	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-pre-upgrade	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-serde	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-service	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-service-rpc	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-shims	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-spark-client	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-standalone-metastore	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-storage-api	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-streaming	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-testutils	3.1.3000.7.1.1.0-565
	org.apache.hive	hive-vector-code-gen	3.1.3000.7.1.1.0-565
	org.apache.hive	kafka-handler	3.1.3000.7.1.1.0-565
	org.apache.hive.hcatalog	hive-hcatalog-core	3.1.3000.7.1.1.0-565
	org.apache.hive.hcatalog	hive-hcatalog-pig-adapter	3.1.3000.7.1.1.0-565
	org.apache.hive.hcatalog	hive-hcatalog-server-extensions	3.1.3000.7.1.1.0-565
	org.apache.hive.hcatalog	hive-hcatalog-streaming	3.1.3000.7.1.1.0-565
	org.apache.hive.hcatalog	hive-webhcat	3.1.3000.7.1.1.0-565
	org.apache.hive.hcatalog	hive-webhcat-java-client	3.1.3000.7.1.1.0-565
	org.apache.hive.shims	hive-shims-0.20	3.1.3000.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.hive.shims	hive-shims-0.23	3.1.3000.7.1.1.0-565
	org.apache.hive.shims	hive-shims-common	3.1.3000.7.1.1.0-565
	org.apache.hive.shims	hive-shims-scheduler	3.1.3000.7.1.1.0-565
Apache Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector_2.11	1.0.0.7.1.1.0-565
Apache Kafka	org.apache.kafka	connect	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-api	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-basic-auth-extension	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-file	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-json	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-mirror	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-mirror-client	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-runtime	2.4.1.7.1.1.0-565
	org.apache.kafka	connect-transforms	2.4.1.7.1.1.0-565
	org.apache.kafka	generator	2.4.1.7.1.1.0-565
	org.apache.kafka	jmh-benchmarks	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-clients	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-examples	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-log4j-appender	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-examples	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-scala_2.11	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-scala_2.12	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-scala_2.13	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-test-utils	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-0100	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-0101	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-0102	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-0110	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-10	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-11	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-20	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-21	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-22	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-streams-upgrade-system-tests-23	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka-tools	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka_2.11	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka_2.12	2.4.1.7.1.1.0-565
	org.apache.kafka	kafka_2.13	2.4.1.7.1.1.0-565

Project	groupId	artifactId	version
Apache Knox	org.apache.knox	gateway-adapter	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-admin-ui	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-applications	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-cloud-bindings	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-cm-integration	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-demo-ldap	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-demo-ldap-launcher	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-discovery-ambari	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-discovery-cm	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-docker	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-i18n	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-i18n-logging-log4j	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-i18n-logging-sl4j	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-ha	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-identity-assertion-common	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-identity-assertion-concat	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-identity-assertion-hadoop-groups	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-identity-assertion-pseudo	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-identity-assertion-regex	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-identity-assertion-switchcase	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-jersey	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-rewrite	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-rewrite-common	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-rewrite-func-hostmap-static	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-rewrite-func-inbound-query-param	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-rewrite-func-service-registry	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-rewrite-step-encrypt-uri	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-rewrite-step-secure-query	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-authc-anon	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-authz-acls	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-authz-composite	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-clientcert	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-hadoopauth	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-jwt	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-pac4j	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-preauth	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-shiro	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-provider-security-webappsec	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-release	1.3.0.7.1.1.0-565



Project	groupId	artifactId	version
	org.apache.knox	gateway-server	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-server-launcher	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-server-xforwarded-filter	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-admin	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-as	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-definitions	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-hashicorp-vault	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-hbase	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-health	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-hive	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-idbroker	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-impala	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-jkg	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-knoxsso	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-knoxssout	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-knoxtoken	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-livy	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-metadata	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-nifi	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-nifi-registry	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-remoteconfig	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-rm	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-storm	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-test	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-tgs	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-vault	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-service-webhdfs	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-shell	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-shell-launcher	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-shell-release	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-shell-samples	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-spi	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-test	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-test-idbroker	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-test-release-utils	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-test-utils	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-topology-simple	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-util-common	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-util-configinjector	1.3.0.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.knox	gateway-util-launcher	1.3.0.7.1.1.0-565
	org.apache.knox	gateway-util-urltemplate	1.3.0.7.1.1.0-565
	org.apache.knox	hadoop-examples	1.3.0.7.1.1.0-565
	org.apache.knox	knox-cli-launcher	1.3.0.7.1.1.0-565
	org.apache.knox	knox-homepage-ui	1.3.0.7.1.1.0-565
	org.apache.knox	webhdfs-kerb-test	1.3.0.7.1.1.0-565
	org.apache.knox	webhdfs-test	1.3.0.7.1.1.0-565
Apache Kudu	org.apache.kudu	kudu-backup-tools	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-backup2_2.11	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-client	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-client-tools	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-hive	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-mapreduce	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-spark2-tools_2.11	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-spark2_2.11	1.12.0.7.1.1.0-565
	org.apache.kudu	kudu-test-utils	1.12.0.7.1.1.0-565
Apache Livy	org.apache.livy	livy-api	0.6.0.7.1.1.0-565
	org.apache.livy	livy-client-common	0.6.0.7.1.1.0-565
	org.apache.livy	livy-client-http	0.6.0.7.1.1.0-565
	org.apache.livy	livy-core_2.11	0.6.0.7.1.1.0-565
	org.apache.livy	livy-examples	0.6.0.7.1.1.0-565
	org.apache.livy	livy-integration-test	0.6.0.7.1.1.0-565
	org.apache.livy	livy-repl_2.11	0.6.0.7.1.1.0-565
	org.apache.livy	livy-rsc	0.6.0.7.1.1.0-565
	org.apache.livy	livy-scala-api_2.11	0.6.0.7.1.1.0-565
	org.apache.livy	livy-server	0.6.0.7.1.1.0-565
	org.apache.livy	livy-test-lib	0.6.0.7.1.1.0-565
	org.apache.livy	livy-thriftserver	0.6.0.7.1.1.0-565
	org.apache.livy	livy-thriftserver-session	0.6.0.7.1.1.0-565
Apache Lucene	org.apache.lucene	lucene-analyzers-common	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-icu	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-kuromoji	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-morfologik	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-nori	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-openslp	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-phonetic	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-smartcn	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-analyzers-stempel	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-backward-codecs	8.4.1.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.lucene	lucene-benchmark	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-classification	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-codecs	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-core	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-demo	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-expressions	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-facet	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-grouping	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-highlighter	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-join	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-memory	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-misc	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-monitor	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-queries	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-queryparser	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-replicator	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-sandbox	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-spatial	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-spatial-extras	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-spatial3d	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-suggest	8.4.1.7.1.1.0-565
	org.apache.lucene	lucene-test-framework	8.4.1.7.1.1.0-565
Apache Oozie	org.apache.oozie	oozie-client	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-core	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-distro	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-examples	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-fluent-job-api	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-fluent-job-client	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-server	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-distcp	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-git	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-hcatalog	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-hive	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-hive2	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-oozie	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-spark	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-sqoop	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-sharelib-streaming	5.1.0.7.1.1.0-565
	org.apache.oozie	oozie-tools	5.1.0.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.oozie	oozie-zookeeper-security-tests	5.1.0.7.1.1.0-565
	org.apache.oozie.test	oozie-mini	5.1.0.7.1.1.0-565
Apache ORC	org.apache.orc	orc-core	1.5.1.7.1.1.0-565
	org.apache.orc	orc-examples	1.5.1.7.1.1.0-565
	org.apache.orc	orc-mapreduce	1.5.1.7.1.1.0-565
	org.apache.orc	orc-shims	1.5.1.7.1.1.0-565
	org.apache.orc	orc-tools	1.5.1.7.1.1.0-565
Apache Phoenix	org.apache.phoenix	phoenix-client	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-core	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-hive	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-load-balancer	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-perf	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-queryserver	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-queryserver-client	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-server	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-spark	5.0.0.7.1.1.0-565
	org.apache.phoenix	phoenix-tracing-webapp	5.0.0.7.1.1.0-565
Apache Ranger	org.apache.ranger	conditions-enrichers	2.0.0.7.1.1.0-565
	org.apache.ranger	credentialbuilder	2.0.0.7.1.1.0-565
	org.apache.ranger	embeddedwebserver	2.0.0.7.1.1.0-565
	org.apache.ranger	jisql	2.0.0.7.1.1.0-565
	org.apache.ranger	ldapconfigcheck	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-atlas-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-atlas-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-distro	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-examples-distro	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-hbase-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-hbase-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-hdfs-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-hdfs-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-hive-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-hive-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-kafka-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-kafka-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-kms	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-kms-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-kms-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-knox-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-knox-plugin-shim	2.0.0.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.ranger	ranger-kudu-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-kylin-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-kylin-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-nifi-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-nifi-registry-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-ozone-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-ozone-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-plugin-classloader	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-plugins-audit	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-plugins-common	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-plugins-cred	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-plugins-installer	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-raz-adls	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-raz-hook-abfs	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-raz-intg	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-raz-processor	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-sampleapp-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-schema-registry-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-solr-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-solr-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-sqoop-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-sqoop-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-storm-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-storm-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-tagsync	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-tools	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-util	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-yarn-plugin	2.0.0.7.1.1.0-565
	org.apache.ranger	ranger-yarn-plugin-shim	2.0.0.7.1.1.0-565
	org.apache.ranger	sampleapp	2.0.0.7.1.1.0-565
	org.apache.ranger	unixauthclient	2.0.0.7.1.1.0-565
	org.apache.ranger	unixauthservice	2.0.0.7.1.1.0-565
	org.apache.ranger	unixusersync	2.0.0.7.1.1.0-565
Apache Solr	org.apache.solr	solr-analysis-extras	8.4.1.7.1.1.0-565
	org.apache.solr	solr-analytics	8.4.1.7.1.1.0-565
	org.apache.solr	solr-cell	8.4.1.7.1.1.0-565
	org.apache.solr	solr-clustering	8.4.1.7.1.1.0-565
	org.apache.solr	solr-core	8.4.1.7.1.1.0-565
	org.apache.solr	solr-dataimporthandler	8.4.1.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.solr	solr-dataimporthandler-extras	8.4.1.7.1.1.0-565
	org.apache.solr	solr-jaegertracer-configurator	8.4.1.7.1.1.0-565
	org.apache.solr	solr-langid	8.4.1.7.1.1.0-565
	org.apache.solr	solr-ltr	8.4.1.7.1.1.0-565
	org.apache.solr	solr-prometheus-exporter	8.4.1.7.1.1.0-565
	org.apache.solr	solr-security-util	8.4.1.7.1.1.0-565
	org.apache.solr	solr-solrj	8.4.1.7.1.1.0-565
	org.apache.solr	solr-test-framework	8.4.1.7.1.1.0-565
	org.apache.solr	solr-velocity	8.4.1.7.1.1.0-565
Apache Spark	org.apache.spark	spark-avro_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-catalyst_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-core_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-graphx_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-hadoop-cloud_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-hive-thriftserver_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-hive_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-kubernetes_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-kvstore_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-launcher_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-mllib-local_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-mllib_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-network-common_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-network-shuffle_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-network-yarn_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-repl_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-sketch_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-sql-kafka-0-10_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-sql_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-streaming-kafka-0-10_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-streaming-kafka-0-8-assembly_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-streaming-kafka-0-8_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-streaming_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-tags_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-unsafe_2.11	2.4.0.7.1.1.0-565
	org.apache.spark	spark-yarn_2.11	2.4.0.7.1.1.0-565
Apache Sqoop	org.apache.sqoop	sqoop	1.4.7.7.1.1.0-565
	org.apache.sqoop	sqoop-test	1.4.7.7.1.1.0-565
Apache Tez	org.apache.tez	hadoop-shim	0.9.1.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.tez	hadoop-shim-2.8	0.9.1.7.1.1.0-565
	org.apache.tez	tez-api	0.9.1.7.1.1.0-565
	org.apache.tez	tez-aux-services	0.9.1.7.1.1.0-565
	org.apache.tez	tez-common	0.9.1.7.1.1.0-565
	org.apache.tez	tez-dag	0.9.1.7.1.1.0-565
	org.apache.tez	tez-examples	0.9.1.7.1.1.0-565
	org.apache.tez	tez-ext-service-tests	0.9.1.7.1.1.0-565
	org.apache.tez	tez-history-parser	0.9.1.7.1.1.0-565
	org.apache.tez	tez-javadoc-tools	0.9.1.7.1.1.0-565
	org.apache.tez	tez-job-analyzer	0.9.1.7.1.1.0-565
	org.apache.tez	tez-mapreduce	0.9.1.7.1.1.0-565
	org.apache.tez	tez-protobuf-history-plugin	0.9.1.7.1.1.0-565
	org.apache.tez	tez-runtime-internals	0.9.1.7.1.1.0-565
	org.apache.tez	tez-runtime-library	0.9.1.7.1.1.0-565
	org.apache.tez	tez-tests	0.9.1.7.1.1.0-565
	org.apache.tez	tez-yarn-timeline-cache-plugin	0.9.1.7.1.1.0-565
	org.apache.tez	tez-yarn-timeline-history	0.9.1.7.1.1.0-565
	org.apache.tez	tez-yarn-timeline-history-with-acls	0.9.1.7.1.1.0-565
	org.apache.tez	tez-yarn-timeline-history-with-fs	0.9.1.7.1.1.0-565
Apache Zeppelin	org.apache.zeppelin	sap	0.8.2.7.1.1.0-565
	org.apache.zeppelin	spark-interpreter	0.8.2.7.1.1.0-565
	org.apache.zeppelin	spark-scala-2.11	0.8.2.7.1.1.0-565
	org.apache.zeppelin	spark-shims	0.8.2.7.1.1.0-565
	org.apache.zeppelin	spark2-shims	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-alluxio	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-angular	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-bigquery	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-cassandra_2.10	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-display	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-elasticsearch	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-file	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-flink_2.10	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-groovy	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-hbase	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-ignite_2.10	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-interpreter	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-jdbc	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-jupyter	0.8.2.7.1.1.0-565
	org.apache.zeppelin	zeppelin-kylin	0.8.2.7.1.1.0-565

Project	groupId	artifactId	version
	org.apache.zookeeper	zookeeper-lens	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-livy	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-markdown	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-neo4j	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-pig	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-python	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-scio_2.10	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-server	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-shell	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-spark-dependencies	0.8.2.7.1.1.0-565
	org.apache.zookeeper	zookeeper-zengine	0.8.2.7.1.1.0-565
Apache ZooKeeper	org.apache.zookeeper	zookeeper	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-client-c	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-contrib-loggraph	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-contrib-rest	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-docs	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-jute	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-recipes-election	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-recipes-lock	3.5.5.7.1.1.0-565
	org.apache.zookeeper	zookeeper-recipes-queue	3.5.5.7.1.1.0-565

## What's New In Cloudera Runtime 7.1.1

This version of Cloudera Runtime provides you with several new capabilities. Learn how the new features and improvements benefit you.

### What's New in Apache Atlas

This topic lists new features for Apache Atlas in this release of Cloudera Runtime.

#### Cloudera Navigator Data Management upgrade to Apache Atlas

In CDP, Apache Atlas fulfills the metadata collection role that in CDH was filled by Cloudera Navigator Data Management. The upgrade to CDP provides a method to migrate Navigator content, including technical and business metadata, to Atlas. For information, see [Migrating Navigator content to Atlas](#). For the cluster audit functionality handled by Navigator, see the (production version of) access auditing provided by Apache Ranger see [Ranger Audit Overview](#).

#### Business Metadata: Entity model extensions

This release of Atlas provides the ability for data stewards to add custom attributes to existing entity types and set their values on existing entities. This functionality allows an organization to extend its enterprise data model with



curated master data attributes that have specific meaning for the business. Business Metadata attributes are defined centrally and can be used on designated entity types. Administrators can control who can view, add values to, and create or update set collections of Business Metadata attributes. Privileged users can add free-form values or select from predefined values to the attribute for a given entity. For more information, see [Leveraging Business Metadata](#).

### Bulk import of Business Metadata attribute associations

Atlas provides an interface to import a list of assignments of Business Metadata attributes to entities. The list includes information to uniquely identify the Business Metadata attribute and the targeted entity. The list can be formatted as comma-separated values (.CSV) or Microsoft Excel (.XLS) formatted file. For more information, see [Importing Business Metadata associations in bulk](#)

### Bulk import of Glossary terms

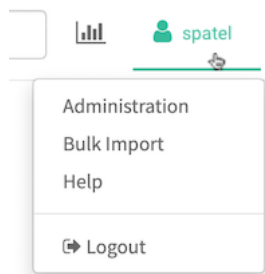
Atlas provides an interface to import a list of terms into existing Glossaries. The list can include any or all of the metadata associated with a given term. The list can be formatted as comma-separated values (.CSV) or Microsoft Excel (.XLS) formatted file. For more information, see [Importing Glossary terms in bulk](#)

### Administrator features have a home in the Atlas UI

The Atlas UI now contains an Administration section available to users with administrator privileges:

- Review system-level audits, such as created by entity purge events. See [Auditing purged entities](#).
- Create enumerations for use as attribute values. See [Defining Apache Atlas enumerations](#).
- Create Business Metadata attributes. See [Creating Business Metadata](#).

Open the Administration section from the user menu at the top right of the Atlas UI.



### Purge of deleted entities

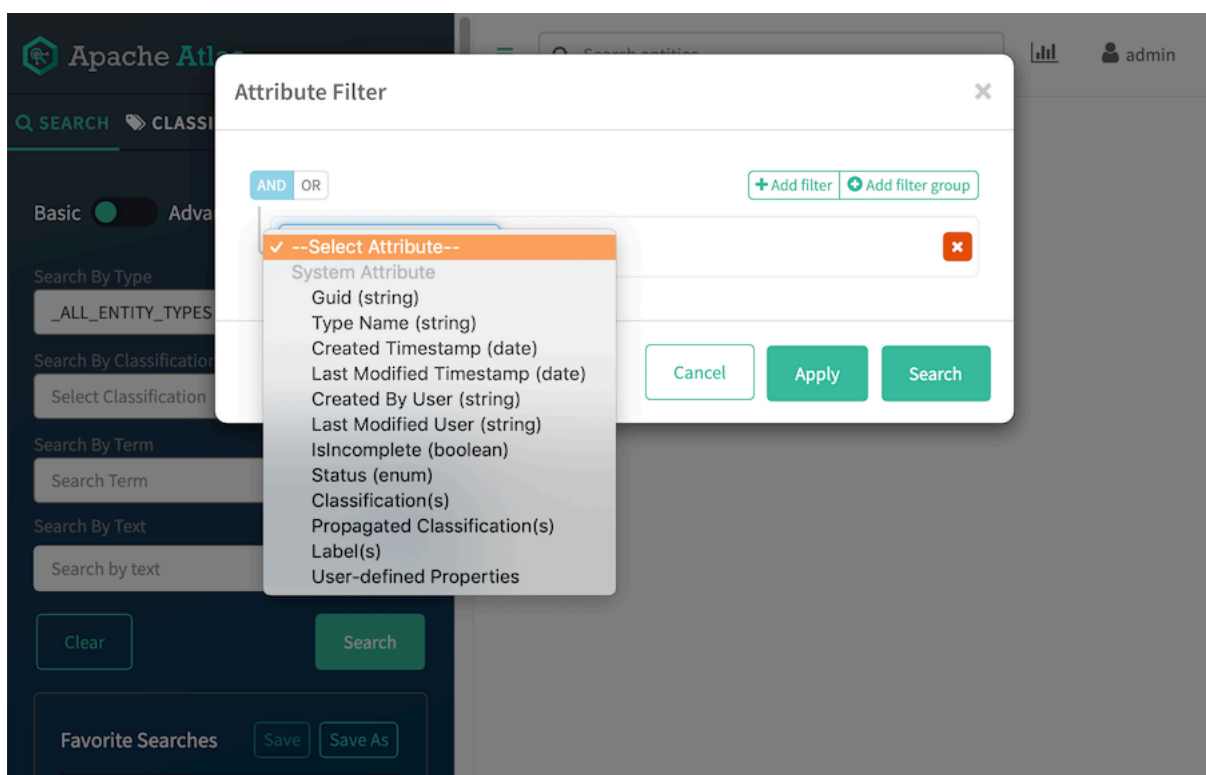
Atlas now provides the ability to clear the metadata for entities that represent data assets and operation that no longer exist on the cluster. The purge functionality is available to users with administrator privilege; run a REST API command that lists one or more GUID values for the deleted entities. For more information, see [Purging deleted entities](#).

### Enhancements to Basic Search in Atlas

Atlas Basic Search includes a filter to allow users to search for entities based on values of entity attributes. In this release, the search filter includes access to system attributes, labels, classifications, and user-defined properties. The filter allows users to build logical combinations of search criteria, including multiple classifications. For more information, see [Using Basic Search](#).

### System attributes filter searches

Atlas basic and advanced search now allow you to filter based on system attribute values, including when and by whom an entity was created. Classifications are also modeled as system attributes, so this change allows you to filter on the classifications assigned to an entity and to distinguish between classifications and propagated classifications. System attributes are available in the search filter.



For information on using system attributes in Advanced Search, see [Apache Atlas metadata attributes](#).

## What's New in Cruise Control

This topic lists new features for Cruise Control in this release of Cloudera Runtime.

### Support for Cruise Control Added

Cruise Control is a Kafka load balancing component that can be used in large Kafka installations. Cruise Control can automatically balance the partitions based on specific conditions, and when adding or removing Kafka brokers.

For more information, see the [Cruise Control documentation](#).

## What's new in DAS

This topic lists new features in DAS in this release of Cloudera Runtime.

- In a CDP Private Cloud Base deployment, you can group admin users and specify a list of admin groups in the `admin_groups` field through the DAS Ambari configuration who need administrative privilege. The admin group feature is not available in Cloudera Manager.
- DAS displays all the DAG IDs that are associated with a particular Hive query, and also displays the DAG graph corresponding to each DAG ID.
- The Hive and Tez configurations are separately displayed on the DAS UI. You can view the Tez-related configurations for a query that has a DAG ID associated with it on the DAG Configurations tab in the DAG Info section on the **Query details** and the **Query Compare** page.
- The Edit button on the **Query Details** page redirects you to the **Compose** page and allows you to edit the selected query in the query composer.

## What's New in Apache HBase

This topic lists new features for Apache HBase in this release of Cloudera Runtime.

If you are upgrading from upgrading from HDP 2.x or CDH 5.x to CDP, you must review the pre-upgrade steps here: [Preparing HBase for upgrade](#), and read about the Apache HBase API changes documented here: [Deprecation Notices in Apache HBase](#).

### HBase-Spark Connector

The HBase-Spark Connector bridges the gap between the simple HBase Key Value store and complex relational SQL queries and enables users to perform complex data analytics on top of HBase using Spark. For more information, see [Using the HBase-Spark connector](#).

### Store Medium Objects (MOBs)

Cloudera's OpDB has a new feature called distributed MOB compaction. This feature overcomes a drawback of the older implementations of MOB compaction by moving maintenance of MOB data files from a centralized process handled by the HBase Master to a parallel process that is distributed across the RegionServers. For more information see, [Storing Medium Objects \(MOBs\)](#).

## What's New in Apache Hadoop HDFS

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

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

## What's New in Apache Hive

This topic lists new Hive features in this release of Cloudera Runtime.

- Scheduled Queries, Rebuilding Materialized Views Automatically Using SQL

You can schedule Hive queries to run on a recurring basis, monitor query progress, temporarily ignore a query schedule, and limit the number running in parallel. You can use scheduled queries to start compaction and periodically rebuild materialized views, for example. For details, see the [Apache Hive Language Manual](#). In CDP, you need to [enable scheduled queries](#).

- Auto-translation for Spark-Hive reads, no HWC session needed

Reads Hive ACID tables in HMS from Spark directly or through HWC based your configuration of spark.sql.extensions. The HWC session is created transparently. Use existing Spark application code without modification.

- [Hive Warehouse Connector Spark direct reads](#)

Spark Direct Reader is a Spark Datasource V1 implementation for reading Hive ACID, transactional tables from Spark. Spark Direct Reader is intended to be used for Extract Transform Load (ETL) or Extract Load Transform (ELT) processes.

- [Authorization of external file writes from Spark](#)

Ranger now authorizes read/write access to external files from Spark through the Hive metastore API (HMS API) in addition to read/write access to managed Hive tables from Spark through HiveServer (HS2).

- [Specifying a top level directory for managed tables when creating a Hive database](#)

## What's New in Hue

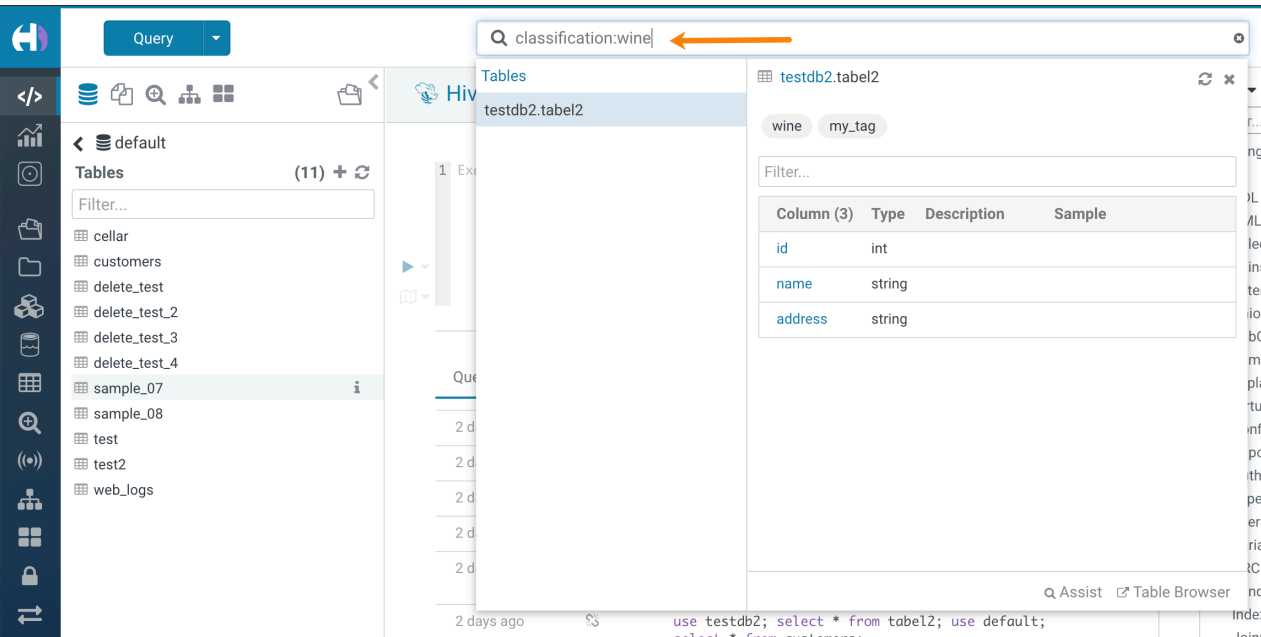
This topic lists new features for Hue in this release of Cloudera Runtime.

Support for Hive on Tez

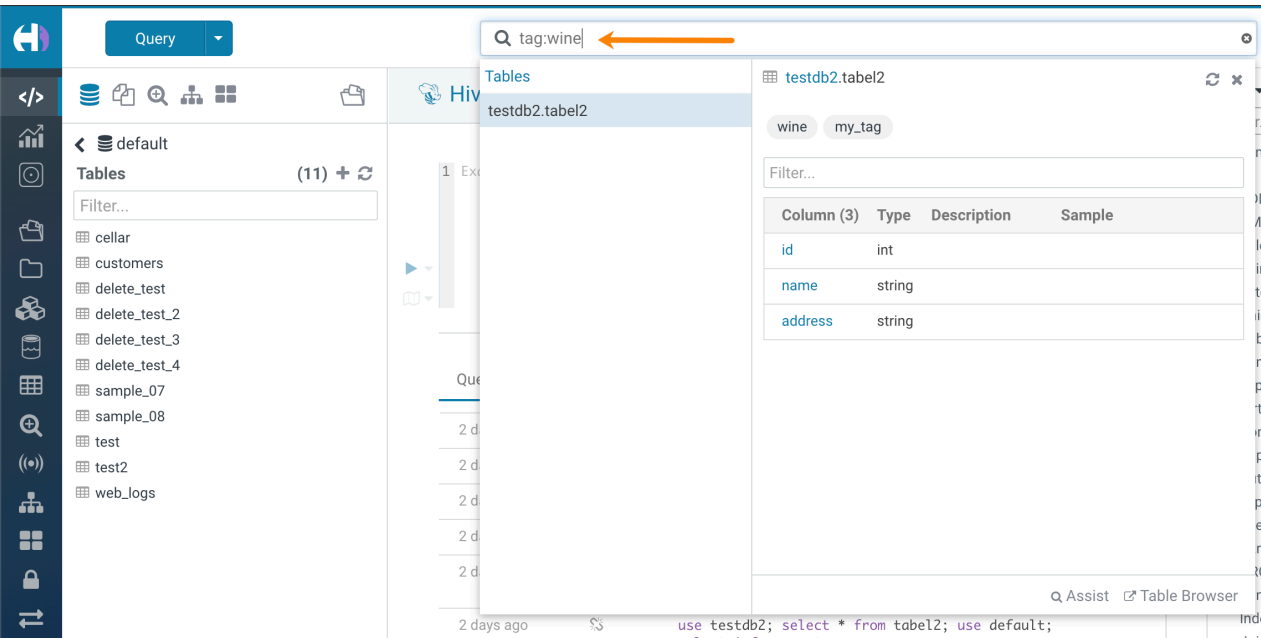
Hue in Cloudera Runtime now supports Hive using Tez as its execution engine.

Integration with Apache Atlas data catalog

You can now locate tables by searching for Atlas classifications in Hue by specifying the classification search term in the search box at the top of the page:

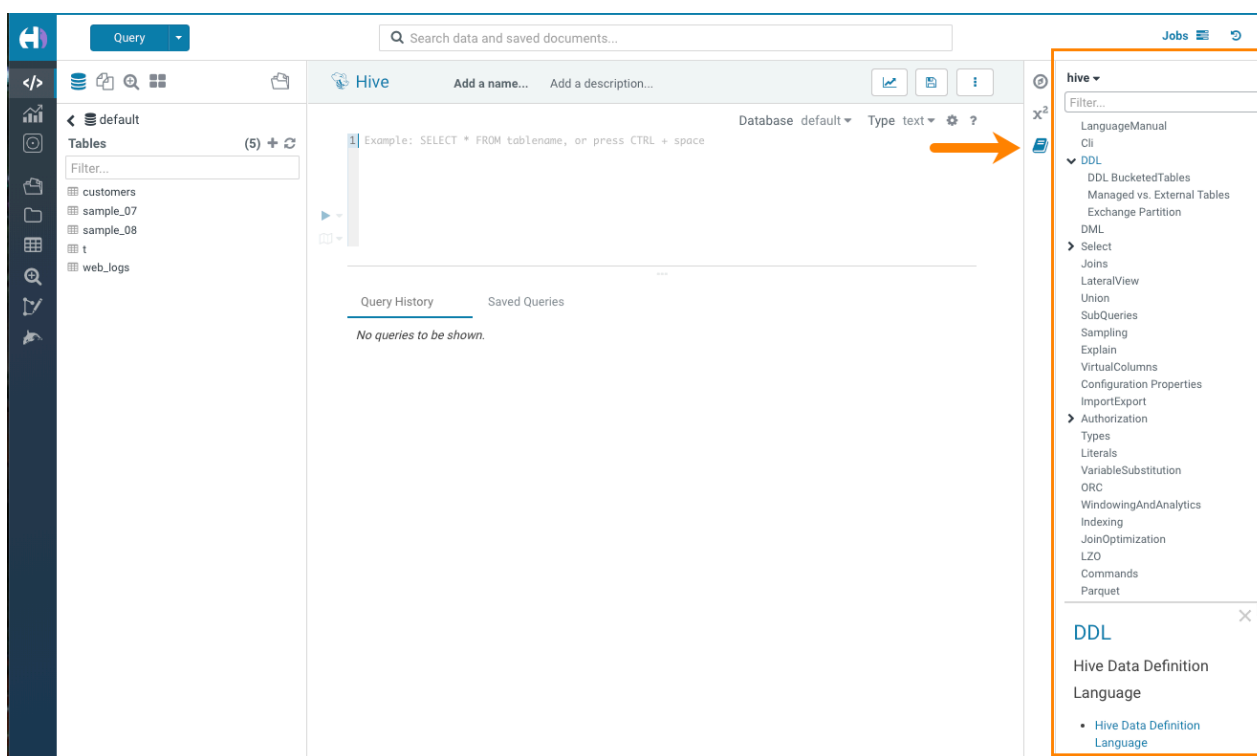


You can also locate tables by searching for Atlas classifications by specifying the tag search term:



Hive Language Reference

A language reference has been added for Hive:



When you select the Hive query editor, click the book icon  to the right of the editor window to launch the Hive language reference.

## What's New in Apache Impala

This topic lists new features for Apache Impala in this release of Cloudera Runtime.

### Optimized performance for multi-threaded query execution

Multi-threaded query execution can be enabled manually on a per-query basis using the `mt_dop` query option for all `SELECT` queries. Previously queries with joins were not supported.

For details, see [MT\\_DOP query option](#).

### Improved read performance for ORC tables with nested type columns

### Improved performance for the automatic updates of metadata

If Impala inserts into a table it refreshes the underlying table/partition. When the configuration `enable_insert_events` is set to `True` Impala will generate `INSERT` event types which when received by other Impala clusters will automatically refresh the tables or partitions. Event processing must be `ON`, for this property to work.

### Started generating Ranger audit logs when column masking policy is applied in a policy

### Reduced the Impala runtime image size and used UBI base image

### Increased scratch capacity

To help reduce spilling to disk:

- Added startup parameter to support Spill-to-disk compression to increase effective scratch capacity by 2.5x.
- Added startup parameter to reclaim space in scratch files.

### Improved data cache performance

- Improved the efficiency of the data cache by providing an option to use a different cache eviction algorithm (LIRS).

### Support for Kudu Date and Varchar column types

### Support for reading ZSTD-compressed text files

For details, see [Using Text Data Files](#).

### Improved read performance of ORC tables

### Improved Impala resiliency

This release adds client retry support in the impala-shell. For details about installing the impala-shell, see [Using Impala shell](#).

### broadcast\_bytes\_limit query option

In this release, you can set a limit for the size of a broadcast input. For details, see [Impala Query Options](#).

### ORC stability and performance improvements

### ORC reads enabled by default

Impala stability and performance have been improved. Consequently, ORC reads are now enabled in Impala by default. To disable, set `-enable_orc_scanner` to false when starting the cluster.

### Constraints

This release adds support for primary and foreign key constraints, but in this release the constraints are advisory and intended for estimating cardinality during query planning in a future release. There is no attempt to enforce constraints. For details, see the “Constraints” section of [Create Table Statement](#).

### Enhanced external Kudu table

By default HMS implicitly translates internal Kudu tables to external Kudu tables with the `'external.table.purge'` property set to true. These tables behave similar to internal tables. You can explicitly create such external Kudu tables. For details, see the “External Kudu Tables” section of [Create Table Statement](#).

### Ranger column masking

This release supports Ranger column masking, which hides sensitive columnar data in Impala query output. For example, you can define a policy that reveals only the first or last four characters of column data. Column masking is enabled by default. For details, see the “Ranger Column Masking” section in [Impala Authorization](#).

## What's New in Apache Kafka

This topic lists new features for Apache Kafka in this release of Cloudera Runtime.

### Rebase on Kafka 2.4.1

Kafka shipped with this version of Cloudera Runtime is based on Apache Kafka 2.4.1. For more information, see [Apache Kafka Notable Changes for versions 2.4.0 and 2.4.1](#), as well as the [Apache Kafka Release Notes for versions 2.4.0 and 2.4.1](#) in the upstream documentation.

### Support for Kafka Connect Added

Support for Kafka Connect is added. In CDP Kafka Connect is implemented in the form of a Kafka service role. The role is called Kafka Connect. In addition, support for Kafka Connect is also added to SMM. Users can from now on manage, monitor, and interact with Kafka Connect either through the SMM UI or SMM REST API. For more information see the [Kafka Connect documentation](#) as well as the [Kafka Connect SMM documentation](#).

### Cloudera developed HDFS and Amazon S3 Sink Connectors Available

Alongside the addition of Kafka Connect support, two Cloudera developed connectors are also added and made available for use. These are the HDFS Sink and Amazon S3 Sink connectors. For more information, see the [Connector documentation](#).

### New Command Line Tool kafka-leader-election

The kafka-preferred-replica-election.sh command line tool has been deprecated in upstream Apache Kafka 2.4.0. It has been replaced by kafka-leader-election.sh. The new tool is available for use in Runtime. In addition, an alternative for it is also provided. The alternative is kafka-leader-election. Both kafka-preferred-replica-election.sh and its alternative, kafka-preferred-replica-election are still available for use, however these are deprecated and will be removed in a future release.

### kafka-sentry command line tool alternative removed

Ranger has replaced Sentry in all versions of CDP Runtime. The kafka-sentry.sh command line tool and its alternative, kafka-sentry, have therefore been deprecated and removed.

### Collection of Producer Metrics is Enabled by Default

The Enable Producer Metrics (producer.metrics.enable) property is now set to true by default, as a result collection of producer metrics is now enabled by default.

### Collection of Partition Level Metrics is Enabled by Default

Cloudera Manager now collects Kafka topic partition level metrics by default. This change is introduced to make the installation of Streams Messaging Manager seamless. The change however, also introduces a limitation. For more information, see the Limitations section in [Kafka Known Issues](#).

## What's New in Apache Knox

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

## What's New in Apache Kudu

This topic lists new features for Apache Kudu in this release of Cloudera Runtime.

### Fine-grained authorization using Ranger

Kudu now supports native fine-grained authorization via integration with Apache Ranger (in addition to integration with Apache Sentry). Kudu may now enforce access control policies defined for Kudu tables and columns stored in Ranger.

### Proxy support using Knox

Kudu's web UI now supports proxying via Apache Knox. Kudu can be deployed in a firewalled state behind a Knox Gateway which will forward HTTP requests and responses between clients and the Kudu web UI.

### Support for HTTP keep-alive

Kudu's web UI now supports HTTP keep-alive. Operations that access multiple URLs will now reuse a single HTTP connection, improving their performance.

### Rolling-restart without stopping on-going Kudu workloads

The kudu tserver quiesce tool is added to quiesce tablet servers. While a tablet server is quiescing, it will stop hosting tablet leaders and stop serving new scan requests. This can be used to orchestrate a rolling restart without stopping on-going Kudu workloads.

### Auto time source support for HybridClock timestamps

Introduced auto time source for HybridClock timestamps. With `--time_source=auto` in AWS and GCE cloud environments, Kudu masters and tablet servers use the built-in NTP client synchronized with dedicated NTP servers available via host-only networks. With `--time_source=auto` in environments other than AWS/GCE, Kudu masters and tablet servers rely on local machine's clock synchronized by NTP. The default setting for the HybridClock time source (`--time_source=system`) is backward-compatible, requiring the local machine's clock to be synchronized by the kernel's NTP discipline.

### Ability to move replicas away from a tablet server

The kudu cluster rebalance tool now supports moving replicas away from specific tablet servers by supplying the `--ignored_tservers` and `--move_replicas_from_ignored_tservers` arguments.

### Ability to specify table creation options using JSON

The kudu table create tool is added to allow users to specify table creation options using JSON.

### Ability to automatically rebalance tablet replicas among tablet servers

An experimental feature is added to Kudu that allows it to automatically rebalance tablet replicas among tablet servers. The background task can be enabled by setting the `--auto_rebalancing_enabled` flag on the Kudu masters. Before starting auto-rebalancing on an existing cluster, the CLI rebalancer tool should be run first.

### Support for DATE and VARCHAR data types

Kudu now supports DATE and VARCHAR data types.

### Optimizations and improvements

- The Write Ahead Log file segments and index chunks are now managed by Kudu's file cache. With that, all the long-lived file descriptors used by Kudu are managed by the file cache, and there's no longer a need for capacity planning file descriptor usage.
- Kudu no longer requires the running of `kudu fs update_dirs` to change a directory configuration or recover from a disk failure
- Kudu tablet servers and masters now expose a tablet-level metric `num_raft_leaders` for the number of tablet replicas hosted on the server
- Kudu's maintenance operation scheduling has been updated to prioritize reducing WAL retention under memory pressure. Kudu would previously prioritize operations that yielded high-memory reduction, which could result in high WAL disk usage in workloads that contained updates
- A new maintenance operation is introduced to remove rowsets that have had all of their rows deleted and whose newest deletes operations are considered ancient



- The built-in NTP client is now fully supported as the time source for Kudu's HybridTime clock. It is no longer marked as experimental. To switch the time source from the existing system time source (which is the default) to the built-in NTP client, use `--time_source=builtin`
- Introduced additional metrics for the built-in NTP client
- Updated /config page of masters' and tablet servers' WebUI to display configured and effective time source.

In addition, the effective list of reference servers for the built-in NTP client is shown there as well, if applicable.

- The processing of Raft consensus vote requests has been improved to be more robust during high contention scenarios like election storms.
- Added a validator to enforce consistency between the maximum size of an RPC and the maximum size of tablet transaction memory, controlled by `--rpc_max_message_size` and `--tablet_transaction_memory` flags correspondingly.

In prior releases, if the limit on the size of RPC requests is increased and the limit on tablet transaction memory size is kept with the default setting, certain Raft transactions could be committed but not applied.

- The metrics endpoint now supports filtering metrics by a metric severity level.
- Many kudu local\_replica tools are updated to not open the block manager, which significantly reduces the amount of IO done when running them
- The Kudu Java client now exposes a way to get the resource metrics associated with a given scanner
- Scan predicates are pushed down to RLE decoders, improving predicate-evaluation-efficiency in some workloads
- The log block manager will now attempt to use multiple threads to open blocks in each data directory, in some tests reducing startup time by up to 20%
- The `raft_term` and `time_since_last_leader_heartbeat` aggregated table metrics will now return the maximum metric reported instead of the sum
- Kudu's tablet server web UI scans page is updated to show the number of round trips per scanner
- Kudu's master and tablet server web UIs are updated to show critical partition information, including tablet count and on-disk size
- Kudu servers now expose the `last_read_elapsed_seconds` and `last_write_elapsed_seconds` tablet-level metrics that indicate how long ago the most recent read and write operations to a given tablet were
- Kudu servers now expose the `transaction_memory_limit_rejections` tablet-level metric that tracks the number of transactions rejected because a given tablet's transactional memory limit was reached

## What's New in Apache Oozie

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

For more information about Oozie, see [Overview of Oozie](#).

## What's New in Apache Hadoop Ozone

This topic lists new features for Apache Hadoop Ozone in this release of Cloudera Runtime..



**Important:** Apache Hadoop Ozone in CDP is available as Beta and is considered to be under development. Do not use this component in your production systems. If you have questions regarding Ozone, contact support by logging a case on the [Cloudera Support Portal](#).

### Support for Recon web user interface

Recon is a centralized monitoring and management service within an Ozone cluster that provides information about the metadata maintained by different Ozone components such as the Ozone Manager (OM) and the Storage Container Manager (SCM). Recon keeps track of the metadata as the cluster is operational, and displays the relevant information through a dashboard and different views on the Recon web user interface. This information helps in understanding the overall state of the Ozone cluster.

For more information, see [Working with Recon web user interface](#)

### Support for High Availability of Ozone Manager

Configuring High Availability (HA) for the Ozone Manager (OM) enables you to run redundant Ozone Managers on your Ozone cluster and prevents the occurrence of a single point of failure in the cluster from the perspective of namespace management.

For more information, see [Overview of Ozone Manager in High Availability](#)

## What's New in Apache Phoenix

This topic lists new features for Apache Phoenix in this release of Cloudera Runtime.

### Apache Phoenix is now generally available

Apache Phoenix in CDP is now generally available from this version of the Cloudera Runtime.

### Connect to PQS through Apache Knox

You can connect to the PQS using the JDBC thin client and the Apache Knox gateway. Apache Knox requires your thin client connection to be over HTTPS. For more information, see [Connect to PQS through Apache Knox](#).

### Manage Apache Phoenix Security using Apache Ranger

Apache Ranger now manages authorization and access control through a user interface that ensures consistent policy administration for both Apache Phoenix and Apache HBase. For more information see [Managing Apache Phoenix Security](#).

For more information about Apache Phoenix, see [Phoenix Overview](#).

### Apache Phoenix-Spark Connector

You can use Apache Phoenix-Spark connector on your secured clusters to perform READ and WRITE operations. The Phoenix-Spark connector allows Spark to load Phoenix tables as Resilient Distributed Datasets (RDDs) or DataFrames, and lets you save them back to Phoenix. For more information see [Understanding Apache Phoenix-Spark Connector](#).

### Apache Phoenix-Hive Connector

You can use Apache Phoenix-Hive connector to access the Phoenix data from Hive without any data transfer. When you use this connector, the Business Intelligence (BI) logic in Hive can access the operational data available in Phoenix. For more information see [Understanding Apache Phoenix-Hive Connector](#).

## What's New in Schema Registry

This topic lists new features for Schema Registry in this release of Cloudera Runtime.

### Schema Registry Distributed with and Included in Runtime

Starting with this release of Cloudera Runtime, Schema Registry is distributed with and included in Runtime. In order to deploy the service you do not need to acquire or install parcels and CSD files separately, all required artifacts now come bundled with Runtime. The service will be by default selectable for deployment when installing a new cluster or when adding new services to a cluster.

## What's New in Cloudera Search

This topic lists new features for Cloudera Search in this release of Cloudera Runtime.

- Cloudera Search shipped with this version of Cloudera Runtime is based on Apache Solr 8.4.1. For more information, see [Major Changes in Solr 8](#) and [Apache Solr Release Notes](#) in the upstream documentation.
- [Logging Slow Queries](#) is now supported.

For more information about Cloudera Search, see [Cloudera Search Overview](#).

### Breaking Changes

- As part of the Solr 8.4 rebase, Cloudera Search is switching to the upstream version of log4j2
- In ParseDateFieldUpdateProcessorFactory the date pattern has changed: Typically a change from uppercase Z to lowercase z is required.

## What's New in Apache Spark

This topic lists new features for Apache Spark in this release of Cloudera Runtime.

### Apache Spark version support

Spark included in Cloudera Runtime versions 7.1.1 for CDP Private Cloud Base is based on Apache Spark version 2.4.5 and contains all the feature content of that release.

### Dynamic Partition Pruning

Added support for dynamic partition pruning. It is disabled by default.

For more information, see [SPARK-11150](#).

## What's New in Sqoop

There are no new features for Sqoop in Cloudera Runtime 7.1.1.

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

This topic lists new features for Streams Replication Manager in this release of Cloudera Runtime.

### Streams Replication Manager Distributed with and Included in Runtime

Starting with this release of Cloudera Runtime, Streams Replication Manager (SRM) is distributed with and included in Runtime. In order to deploy the service you do not need to acquire or install parcels and CSD files separately, all required artifacts now come bundled with Runtime. The service will be by default selectable for deployment when installing a new cluster or when adding new services to a cluster.

## What's new in Streams Messaging Manager

This topic lists new features for Streams Messaging Manager in this release of Cloudera Runtime.

### Streams Messaging Manager Distributed with and Included in Runtime

Starting with this release of Cloudera Runtime, Streams Messaging Manager (SMM) is distributed with and included in Runtime. In order to deploy the service you do not need to acquire or install parcels and CSD files separately. All required artifacts now come bundled with Runtime. The service will be selectable for deployment when installing a new cluster or when adding new services to a cluster.

### Additional new Streams Messaging Manager features

#### Refresh option added on pages of the SMM UI

A refresh button is added on all pages of the SMM UI where a time-range can be set. The refresh button allows you to refresh data displayed on a page without having to refresh from the browser. This enables you to refresh the data displayed without losing filter settings or selections made on a page.

#### Default Value of Streams Messaging Manager Configuration Directory changed

The default value of the Streams Messaging Manager Configuration Directory (`streams.messaging.manager.working.directory`) property, which sets the working directory of SMM, is changed from `var/lib/streamsmgsmgr` to `/var/lib/streams_messaging_manager`.

#### Cloudera Manager Service Monitor Host property added

A new property, Cloudera Manager Service Monitor Host (`cm.metrics.service.monitor.host`), has been added to SMM. This property allows you to configure the host of the CM Service Monitor that SMM connects to when CM Server and CM Service Monitor are installed on different hosts.

If CM Server and CM Service Monitor are installed on different hosts, you must configure the Cloudera Manager Service Monitor Host and Cloudera Manager Service Monitor Port properties to enable SMM to collect metrics.

#### Kafka Connect in SMM

You can use the Kafka Connect option in SMM to manage Kafka Connectors in your cluster.

## What's New in Apache Hadoop YARN

This topic lists new features for Apache Hadoop YARN in this release of Cloudera Runtime.

### Placement Rules

Placement Rules allow you to specify a set of rules for assigning jobs and applications to queues. You can now define placement rules to dynamically create queues based on the rules and the applications would be submitted to those queues. These predefined rules enable you to submit jobs without specifying the queue name at the time of job submission.

For more information about configuring placement rules, see [Configure placement rules](#).

### Data Locality

Capacity Scheduler leverages Delay Scheduling to honor task locality constraints. The scheduler counts the number of missed opportunities when the locality cannot be satisfied and waits for this count to reach a threshold before

relaxing the locality constraint to the next level. You can configure this threshold using the YARN Queue Manager UI.

For more information about configuring data locality, see [Configure data locality](#).

### FPGA as a resource type

You can use FPGA as a resource type. For more information, see [Use FPGA scheduling](#).

### NodeManager Heartbeat

You can now control how many containers can be allocated in each NodeManager heartbeat. You can set the container assignments and off-switch assignments per NodeManager heartbeat.

For more information about configuring NodeManager heartbeat, see [Configure NodeManager heartbeat](#).

### Preemption

Preemption allows applications of higher priority to preempt applications of lower priority. If preemption is enabled, applications of higher priority do not have to wait because applications of lower priority have taken up the available capacity.

For more information on configuring preemption, see [Configure preemption](#).

### Same queue name under different hierarchies

Capacity Scheduler now support queues with the same name under different hierarchies, such as: root.a.q1 and root.b.q2.

### Moving applications between queues (Technical Preview)

Capacity Scheduler now supports moving of application between queues.



**Note:** This feature is a technical preview and considered under development. Do not use this in your production environment. If you have feedback, contact Support by logging a case on our [Cloudera Support Portal](#). Technical preview features are not guaranteed troubleshooting and fixes.

### More reading

For more information about Apache Hadoop YARN, see [YARN Overview](#).

## What's New in Apache ZooKeeper

This topic lists new features for Apache ZooKeeper in this release of Cloudera Runtime.

### ZooKeeper TLS/SSL

TLS/SSL encryption between the ZooKeeper client and the ZooKeeper server and within the ZooKeeper Quorum is supported. TLS/SSL encryption is automatically enabled when AutoTLS is enabled, or you can enable it manually using Cloudera Manager. ClientSSL requires you to enable and configure it by other components that you want to use this feature with. For more information, see [Configure ZooKeeper TLS/SSL using Cloudera Manager](#).

## Deprecation Notices In Cloudera Runtime 7.1.1

Components and features that will be deprecated or removed in this release or a future release.

## 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 HBase

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.

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&lt;E&gt;</code> to <code>org.apache.hadoop.hbase.util.SortedList&lt;E&gt;</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&lt;MasterCoprocessorEnvironment&gt;, 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&lt;MasterCoprocessorEnvironment&gt;, 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&lt;MasterCoprocessorEnvironment&gt;, 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&lt;MasterCoprocessorEnvironment&gt;, List&lt;HTableDescriptor&gt; )</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>voidpostModifyTable ( ObserverContext&lt;MasterCoprocessorEnvironment&gt;, 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&lt;MasterCoprocessorEnvironment&gt;, 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&lt;MasterCoprocessorEnvironment&gt;, 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&lt;MasterCoprocessorEnvironment&gt;, 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&lt;MasterCoprocessorEnvironment&gt;, HTableDescriptor, HRegionInfo[] )</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.



Abstract method voidpreDeleteSnapshot ( ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpreGetTableDescriptors ( ObserverContext<MasterCoproprocessorEnvironment>, List<TableName>, List<HTableDescriptor> ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpreModifyTable ( ObserverContext<MasterCoproprocessorEnvironment>, TableName, HTableDescriptor ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpreRestoreSnapshot ( ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpreSnapshot ( ObserverContext<MasterCoproprocessorEnvironment>, HBaseProtos.SnapshotDescription, HTableDescriptor ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.

## RegionObserver interface changes

The following changes are introduced to the RegionObserver interface.

[#] interface RegionObserver (13)

Change	Result
Abstract method voidpostCloseRegionOperation ( ObserverContext<RegionCoproprocessorEnvironment>, HRegion.Operation ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpostCompactSelection ( ObserverContext<RegionCoproprocessorEnvironment>, Store, ImmutableList<StoreFile> ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpostCompactSelection ( ObserverContext<RegionCoproprocessorEnvironment>, Store, ImmutableList<StoreFile>, CompactionRequest ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpostGetClosestRowBefore ( ObserverContext<RegionCoproprocessorEnvironment>, byte[ ], byte[ ], Result ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method DeleteTrackerpostInstantiateDeleteTracker ( ObserverContext<RegionCoproprocessorEnvironment>, DeleteTracker ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpostSplit ( ObserverContext<RegionCoproprocessorEnvironment>, HRegion, HRegion ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpostStartRegionOperation ( ObserverContext<RegionCoproprocessorEnvironment>, HRegion.Operation ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
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 NoSuchMethodErrorException.
Abstract method voidpostWALRestore ( ObserverContext<RegionCoproprocessorEnvironment>, HRegionInfo, HLogKey, WALEdit ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method InternalScannerpreFlushScannerOpen ( ObserverContext<RegionCoproprocessorEnvironment>, Store, KeyValueScanner, InternalScanner ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.
Abstract method voidpreGetClosestRowBefore ( ObserverContext<RegionCoproprocessorEnvironment>, byte[ ], byte[ ], Result ) has been removed from this interface.	A client program may be interrupted by NoSuchMethodErrorException.



Abstract method <code>StoreFile.Reader.preStoreFileReaderOpen ( ObserverContext&lt;RegionCoproprocessorEnvironment&gt;, FileSystem, Path, FSDataInputStreamWrapper, long, CacheConfig, Reference, StoreFile.Reader)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>void.preWALRestore ( ObserverContext&lt;RegionCoproprocessorEnvironment&gt;, HRegionInfo, HLogKey, WALEdit)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.

### WALObserver interface changes

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

[#] interface `WALObserver`

Change	Result
Abstract method <code>void.postWALWrite ( ObserverContext&lt;WALCoproprocessorEnvironment&gt;, HRegionInfo, HLogKey, WALEdit)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.
Abstract method <code>boolean.preWALWrite ( ObserverContext&lt;WALCoproprocessorEnvironment&gt;, HRegionInfo, HLogKey, WALEdit)</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> exception.

### Scheduler changes

Following methods are now changed to abstract:

[#]class `RpcScheduler` (1)

Change	Result
Abstract method <code>void.dispatch ( CallRunner )</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchMethodError</code> 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 <code>void</code> to <code>boolean</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.

The following abstract methods have been removed:

[#]interface `PriorityFunction` (2)

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

### Server API changes

[#] class `RpcServer` (12)

Change	Result
Type of field <code>CurCall</code> has been changed from <code>java.lang.ThreadLocal&lt;RpcServer.Call&gt;</code> to <code>java.lang.ThreadLocal&lt;RpcCall&gt;</code> .	A client program may be interrupted by <code>NoSuchFieldError</code> exception.

Abstract method <code>int getNumOpenConnections ( )</code> has been added to this class.	This class became abstract and a client program may be interrupted by <code>InstantiationException</code> exception.
Field <code>callQueueSize</code> of type <code>org.apache.hadoop.hbase.util.Counter</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>connectionList</code> of type <code>java.util.List&lt;RpcServer.Connection&gt;</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>maxIdleTime</code> of type <code>int</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>numConnections</code> of type <code>int</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>port</code> of type <code>int</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>purgeTimeout</code> of type <code>long</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>responder</code> of type <code>RpcServer.Responder</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>socketSendBufferSize</code> of type <code>int</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.
Field <code>thresholdIdleConnections</code> of type <code>int</code> has been removed from this class.	A client program may be interrupted by <code>NoSuchFieldError</code> exception.

Following abstract methods are removed:

Change	Result
Abstract method <code>Pair&lt;Message,CellScanner&gt;call ( BlockingService, Descriptors.MethodDescriptor, Message, CellScanner, long, MonitoredRPCHandler )</code> has been removed from this interface.	A client program may be interrupted by <code>NoSuchMethodError</code> 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 <code>clusterIds</code> has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field <code>compressionContext</code> has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field <code>encodedRegionName</code> has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field <code>tablename</code> has been changed from protected to private.	A client program may be interrupted by <code>IllegalAccessError</code> exception.
Access level of field <code>writeTime</code> 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 <code>LOG</code> 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 <code>VERSION</code> 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 RelocationAdmin, 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 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_USEMULTIof 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 IncompatibleClassChangeError or InstantiationException 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 java.util.Map<java.lang.String,master.RegionState> to java.util.List<master.RegionState>.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchMethodError 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 org.apache.hadoop.hbase.protobuf.generated.HBaseProtos.SnapshotDescription to org.apache.hadoop.hbase.client.SnapshotDescription.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchElementException.

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 org.apache.hadoop.hbase.protobuf.generated.ClientProtos.RegionLoadStats to RegionLoadStats.	This method has been removed because the return type is part of the method signature. A client program may be interrupted by NoSuchElementException.

## 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 NoSuchElementException.

## Deprecation notices in Apache Kudu

This topic lists the features and functionality in Apache Kudu that will be deprecated or removed in this release or 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 in Cloudera Search

Features and functionality that will be deprecated or removed in this release or a future release.

### Deprecated

#### Deprecated environment variables and Java options

- Environment variables: SOLR\_CIPHERS\_CONFIG , SOLR\_TLS\_PROTOCOLS\_CONFIG
- and Java options: solr.jetty.ciphers , solr.jetty.tls.protocols



are now deprecated and throw a warning when used. Use the following environment variables/Java Options instead:

Environment Variable	Java Option	Default
SOLR_SSL_PROTOCOLS_INCLUDE	solr.jetty.ssl.protocols.include (formerly solr.jetty.tls.protocols)	TLSv1,TLSv1.1,TLSv1.2,TLSv1.3 *TLSv1.3 will work only if Solr is running on Java11
SOLR_SSL_PROTOCOLS_EXCLUDE	solr.jetty.ssl.protocols.exclude	SSL,SSLv2,SSLv2Hello,SSLv3
SOLR_SSL_CIPHERS_INCLUDE	solr.jetty.ssl.ciphers.include (formerly solr.jetty.ciphers)	^.*\$
SOLR_SSL_CIPHERS_EXCLUDE	solr.jetty.ssl.ciphers.exclude	^.*_(MD5 SHA SHA1)\$,^TLS_RSA_.*\$,^SSL_.*\$,^.*_NULL_.*\$,^.*_anon_.*\$

If both legacy and current configuration values are present, the current configuration overrides the legacy one. Exclude values override include ones.

Deprecated in: Cloudera Search versions shipped with Solr 8.4.1 or later.

Will be removed in: a future release.

### Removed

- Index-time boosts have been removed from Lucene, and are no longer available from Solr. In this release of Cloudera Search, using syntax: { "id": "1", "val\_s": { "value": "foo", "boost": 2.0 } } will throw an error message.
- StandardFilter has been removed from Solr. As this was a non-operator, its removal should cause no issues.

## Deprecation Notices for Apache Kafka

Features and functionality that will be deprecated or removed in this release or 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.

### Removed

#### kafka-sentry

The kafka-sentry.sh command line tool and its alternative, kafka-sentry, have been removed.

## Behavioral Changes In Cloudera Runtime 7.1.1

Behavioral changes denote a marked change in behavior from the previously released version to this version of Cloudera Runtime.

## Behavioral Changes in Cloudera Search

Behavioral changes denote a marked change in behavior from the previously released version to this version of Cloudera Search.

**Cloudera ID:**

CDPD-6731

**Apache component**

Solr

**Apache JIRA:**

SOLR-12167

**Summary:**

Invalid Atomic Update operations now fail

**Details:**

**Previous behavior:**

Invalid Atomic Updates threw a warning message.

**New behavior:**

Invalid Atomic Updates fail with an Exception.

**Cloudera ID**

DOCS-5721

**Apache component**

Solr

**Apache JIRA:**

N/A

**Summary:**

Admin API address has changed

**Details:**

**Previous behavior:**

In Solr 7 both `curl -k --negotiate -u: "https://^hostname -f :8985/solr/?op=GETDELEGATIONTOKEN"` and `curl -k --negotiate -u: "https://^hostname -f :8985/solr/admin?op=GETDELEGATIONTOKEN"` commands worked.

**New behavior:**

In Solr 8 only `curl -k --negotiate -u: "https://^hostname -f :8985/solr/admin?op=GETDELEGATIONTOKEN"` command (with the 'admin' string added) works.

## Behavioral Changes in Apache Hive

Behavioral changes denote a marked change in behavior from the previously released version to this version of Apache Hive.

**Cloudera ID:**

CDPD-12108

**Apache component**

Hive

**Summary:**

Query results are returned unordered

**Details:****Previous behavior:**

When using the Hive Warehouse Connector in JDBC cluster mode, `SELECT ... ORDER BY` returned ordered results.

**New behavior:**

When using the Hive Warehouse Connector in JDBC cluster mode, `SELECT ... ORDER BY` returns unordered results. There is no guarantee on ordering when a SQL query is executed via `"hive.sql()"`. Since the order by action is inside the query sent to HiveServer, and therefore not known to Spark, the data in the relation is not ordered before being returned as a result.

## Behavioral Changes in Apache Hadoop YARN

Behavioral changes denote a marked change in behavior from the previously released version to this version of Apache Hadoop YARN.

**Summary:**

Capacity Scheduler uses the `DominantResourceCalculator` by default.

**Details:****Previous behavior:**

Capacity Scheduler uses the `DefaultResourceCalculator` by default.

**New behavior:**

Capacity Scheduler uses the `DominantResourceCalculator` by default.

## Fixed Issues In Cloudera Runtime 7.1.1

Fixed issues represent issues reported by Cloudera customers that are addressed in this release.

### Fixed issues in Atlas

This section lists the issues that have been fixed since the previous version. See also [What's New in Apache Atlas](#) on page 24.

**CDPD-11434: Lineage graph improvement and lib version updated**

In previous releases, very complicated lineage relations combined with classifications that propagate through the lineage relations could cause the Atlas UI to fail to render the lineage for an entity. This release improves how the Atlas UI manages updating large numbers of lineage relations with classifications.

**CDPD-10399: Upgrade jackson-databind to version 2.9.10.4 for Atlas**

Atlas now uses the Jackson databind library version 2.9.10.4.

**CDPD-9805: CVE: Use correct version of jackson-databind for Impala**

The Atlas Impala bridge has been updated to specify the Jackson databind library in the same way that it is specified in the Hive bridge so that both applications use the same library version.

**CDPD-8240: Atlas - Upgrade to Jetty 9.4.26 to avoid CVEs**

Atlas 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.

**CDPD-8017: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

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

**CDPD-7851: Atlas UI security vulnerabilities**

Atlas includes security headers for all REST API endpoints.

**CDPD-6405: Stored XSS in Atlas**

This release includes a fix for a stored cross-site scripting (XSS) attack vulnerability in the Atlas UI.

**OPSAPS-51224: Atlas custom properties ignored in client services**

When adding a custom property for the atlas-application.properties in Atlas hook-based services such as Hive, HBase, and Impala, the custom property is not reflected in the actual configuration file that Cloudera Manager generates, causing these properties to be ignored.

## Fixed issues in DAS

This section lists the issues that have been fixed since the previous version.

- Fixed several stability issues in the reporting pipeline which caused the report pipeline to get stuck.
- Fixed performance issue in the events pipeline when events are on the S3 filesystem.
- [DWX-2508](#): DAS now shows accurate compilation time for a query on the Query Details Timeline tab.
- [DWX-2220](#): The **Compose** page UI has been enriched with tooltips and also displays the database count.

## Fixed issues in Hadoop

This section lists the issues that have been fixed since the previous version.

**CDPD-9110: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

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

**CDPD-7366: Hadoop - Upgrade to Jetty 9.4.26 to avoid CVEs**

Hadoop 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 HBase

This section lists the issues that have been fixed since the previous version.

**CDPD-7356: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

HBase and HBase connectors has been upgraded to use Guava version 28.1 to avoid CVE-2018-10237.

**CDPD-10091: Upgrade jackson-databind to version 2.9.10.4 to resolve multiple CVEs**

Upgraded Jackson Databind to 2.9.10.4 to address to following CVEs: CVE-2020-9547, CVE-2020-11111, CVE-2020-10672, CVE-2020-10969, CVE-2020-11112, CVE-2020-9548, CVE-2020-9546, CVE-2020-10968, CVE-2020-10673, CVE-2020-11113.

**CDPD-10099: Upgrade Jackson Databind to version 2.10.latest**

Upgrade Jackson to version 2.10.3 to avoid future CVEs.

**CDPD-10532: Update Log4j to address CVE-2019-17571**

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

**CDPD-7724: HBase - Upgrade to Jetty 9.4.26 to avoid CVEs**

HBase 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 HDFS

This section lists the issues that have been fixed since the previous version.

### **CDPD-10091: Upgrade jackson-databind to version 2.9.10.4 to resolve multiple CVEs**

Upgraded Jackson Databind to 2.9.10.4 to address the following CVEs: CVE-2020-9547, CVE-2020-11111, CVE-2020-10672, CVE-2020-10969, CVE-2020-11112, CVE-2020-9548, CVE-2020-9546, CVE-2020-10968, CVE-2020-10673, CVE-2020-11113.

## Fixed Issues in Hive

This section lists the issues that have been fixed since the previous version.

### **CDPD-8015: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

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

### **CDPD-7367: Hive - Upgrade to Jetty 9.4.26 to avoid CVEs**

Hive 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.

### **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`.

## Fixed Issues in Hue

There are no fixed issues in this release of Cloudera Runtime.

## Fixed Issues in Kafka

This section lists the issues that have been fixed since the previous version.

Bug ID: RELENG-8748

Apache JIRA: N/A

Apache Component: Kafka

Summary: Kafka command line tool alternatives unavailable.

Bug ID: CDPD-10143

Apache JIRA: KAFKA-9712

Apache Component: Kafka

Summary: Reflection library update causes ClassLoading issues in Kafka connect (KAFKA-9712 backport).

## Fixed Issues in Impala

There are no fixed issues in this release of Cloudera Runtime.

**CDPD-8014: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

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

**CDPD-7369: Impala - Upgrade to Jetty 9.4.26 to avoid CVEs**

Impala 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 Kudu

This section lists the issues that have been fixed since the previous version.

- **KUDU-2929:** Kudu does not schedule compactions if a server is under memory pressure.
- **KUDU-3036:** DDL operations like ALTER TABLE on tables with huge number of partitions no longer result in a DoS situation for Kudu masters.
- **KUDU-3061:** Fixed a bug where Kudu Java client cannot negotiate a secure connection with Kudu masters and tablet servers if using BouncyCastle JCE provider.
- **KUDU-2904:** Kudu masters will now crash immediately upon hitting a disk failure.
- **KUDU-2992:** Delays in receiving tablet server heartbeats no longer result in an excess amount of RPC traffic between the masters and the tablet servers.
- **KUDU-3008:** Kudu's location placement policy no longer stores all the replicas in a single location when more than one locations are available.
- **KUDU-3035:** The Java client correctly propagates timestamps when sending write batches.
- **KUDU-3099:** The Kudu backup Spark jobs contain a fix so that Kudu no longer returns with a non-zero exit if the job succeeded but backed up no rows.
- **KUDU-3106:** The Kudu Java client can now negotiate a secure connection with Kudu masters and tablet servers when using BouncyCastle.

## Fixed issues in Oozie

This section lists the issues that have been fixed since the previous version.

**CDPD-7357: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

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

**CDPD-7526: Oozie - Upgrade to Jetty 9.4.26 to avoid CVEs**

Oozie 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, and CVE#2019#10247.

**CDPD-10746: Update log4j to address CVE-2019-17571****CDPD-9761: <https://issues.apache.org/jira/browse/OOZIE-3584>: Fork-join action issue arises when action parameter is not resolved.**

There is a sub workflow run in the independent mode that runs a fork action which contains two or more actions. These actions in the fork action run in parallel mode and have a few seconds of 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 workflow state to FAILED. The other actions state which are not yet started will get stuck in PREP state forever. The correct behavior is to KILL the remaining actions as well as the workflow. This issue occurs only when you run this in the 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 uses the spark-hive library from the stack.

**CDPD-9220: <https://issues.apache.org/jira/browse/OOZIE-3586>: Oozie spark actions using --keytab fail due to duplicate distributed cache**

In CDH 6.x, on the Hadoop 3 rebase, it is now a failure if items are added multiple times to the distributed cache. In CDH 5, this was a warning. This is not an issue for most users, as adding multiple times typically is user error, but this completely breaks Spark actions with keytabs (--keytab). 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-7108: <https://issues.apache.org/jira/browse/OOZIE-3561>: Forkjoin validation gets slow when there are more actions in chain.**

In a workflow, if there are more actions, for example, 80 actions one after the other, then the validator code never completes.

**CDPD-7107: <https://issues.apache.org/jira/browse/OOZIE-3551>: Configure working defaults for Spark action in Oozie.**

The following is added to the spark opts section of the spark action:

- --conf spark.yarn.security.tokens.hiveserver2.enabled=false
- --conf spark.yarn.security.tokens.hivestreaming.enabled=false

**CDPD-7106: <https://issues.apache.org/jira/browse/OOZIE-2828>: Query tag is not functional for Hive2 action node in Oozie.**

Query tag is not functional for Hive2 action node in oozie. Workflow is intended to create a hive table using a Hive2 action node. Though workflow runs successfully, the table is not created.

**CDPD-7105: Oozie workflow processing becomes slow after the increasing the rows in WF\_JOBS and WF\_ACTIONS tables.**

Oozie workflow processing becomes slow after the increase of rows in WF\_JOBS and WF\_ACTIONS tables when running against SQL Server.

**CDPD-6877: <https://issues.apache.org/jira/browse/OOZIE-3578>: MapReduce counters cannot be used over 120**

When you create a Map-Reduce action, then it creates more than 120 counters. This displays an exception.

**CDPD-6630: <https://issues.apache.org/jira/browse/OOZIE-3575>: Add credential support for cloud file systems**

Oozie by default gathers delegation tokens for the nodes defined in mapreduce.job.hdfs-servers or oozie.launcher.mapreduce.job.hdfs-servers in case of distcp actions and for the workflow path.

Though this implementation works for HDFS, such implementations are not supported where the job relates resources, which must access runtime and are present on different file systems/buckets and so on.

The following scenarios are addressed: Oozie should obtain delegation token in the following cases:

- The defaultFS is cloud.
- The workload.xml is in cloud.
- Input/output/auxiliary files referred from workflow are in cloud.
- Newly introduced feature - you can define filesystem credentials for the workflow (as its done with Hive or HCAT). This allows you to handle where Oozie is unable to decide the tokens needed at launch time by default and can also get tokens for different cloud storages and buckets too.

**CDPD-5168: <https://issues.apache.org/jira/browse/OOZIE-3381>: Enhance logging of CoordElFunctions.**

Logging enhancements in CoordElFunctions for better supportability.

**CDPD-4826: Oozie TLS does not work with OpenJDK 11**

Oozies web server does not work when TLS is enabled and Open JDK 11 is in use.

## Fixed issues in Ozone

This section lists the issues that have been fixed since the previous version.

### **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

This section lists the issues that have been fixed since the previous version.

### **CDPD-10532: Update Log4j to address CVE-2019-17571**

Replaced Log4j with an internal version to fix CVE-2019-17571.

### **CDPD-11333: Update Netty in Phoenix-Hive connector**

Updated Netty to 4.1.47.Final to fix CVEs in previous versions.

### **CDPD-10452: Exclude ZooKeeper and Guava from Curator in Phoenix**

Harmonize dependencies within CDP.

### **CDPD-10312: Update Scala version for Phoenix**

Scala version updated to 2.11.12. Harmonize dependencies within CDP.

### **CDPD-7371: Phoenix - Upgrade to Jetty 9.4.26 to avoid CVEs**

Phoenix 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 Search

This section lists the issues that have been fixed since the previous version.

### **HBase Lily indexer might fail to write role log files**

In certain scenarios the HBase Lily Indexer (Key-Value Store Indexer) fails to write its role log files.

Workaround: None.

### **Cloudera Bug ID:**

CDH-82599

### **Processing UpdateRequest with delegation token throws NullPointerException**

When using the Spark Crunch Indexer or another client application which utilizes the SolrJ API to send Solr Update requests with delegation token authentication, the server side processing of the request might fail with a NullPointerException.

### **Workaround:**

None.

### **Apache Issue:**

SOLR-13921

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

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



**CDPD-7713: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

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

**CDPD-4731: Upgrade hbase-indexer, Solr, and Search jetty to 9.4.x**

Search 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 Spark

This section lists the issues that have been fixed since the previous version.

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

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

**CDPD-7882: Spark cannot insert into a Hive table with a subset of partition columns dynamic**

This issue is now fixed.

**CDPD-7373: Spark - Upgrade to Jetty 9.4.26 to avoid CVEs**

Spark 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.

**CDPD-7823: Upgrade to Guava 28.1 to avoid CVE-2018-10237**

Spark has been upgraded to use Guava version 28.1 to avoid CVE-2018-10237. Applications that previously depend on old Guava versions (such as guava 11.0.2) may require extensive rewrite to be compatible with Guava 28.1. Applications should ensure they do not introduce older versions of Guava at runtime.

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

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

## Fixed issues in Sqoop

This section lists the issues that have been fixed since the previous version.

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

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

## Fixed Issues in Streams Replication Manager

This section lists the issues that have been fixed since the previous version.

Bug ID: CSP-861

Apache JIRA: N/A

Apache Component: Streams Replication Manager

Summary: The SRM service continues to run after failing to create metrics topic with no error indication in Cloudera Manager.

Bug ID: CSP-876

Apache JIRA: N/A

Apache Component: Streams Replication Manager

Summary: SRM in an unsecured cluster cannot connect to a Kafka in a secure cluster.

Bug ID: CSP-877

Apache JIRA: N/A

Apache Component: Streams Replication Manager

Summary: Cannot configure datahub.sasl.jaas.config in SRM properties.

Bug ID: CSP-1000

Apache JIRA: N/A

Apache Component: Streams Replication Manager

Summary: The srm-control command line tool does not explicitly create the srm-control.<target>.internal topic.

## Fixed Issues in Streams Messaging Manager

This section lists the issues that have been fixed since the previous version.

### **The Condition field in the Alert Policy UI displays incorrect values**

Bug ID: CSP-734

Apache JIRA: N/A

Component: Streams Messaging Manager

Summary: When you edit or update an alert policy with REST API in SMM, the CONDITION field in the Alert Policy UI does not display correct values.

### **Exception in the SMM UI - Not displaying any data**

Bug ID: CSP-875

Apache JIRA: N/A

Component: Streams Messaging Manager

Summary: If the topic (or) producer name contains keywords such as clientId/topic/partition, then SMM would not display the metrics for those resources. This is due to the bug in regular expression which captures the topic (or) partition information from the metric name. SMM also throws "For Input String" error in the UI when there exists another topic whose name is a suffix of the first one. For example, for a topic named with "hello.topic.world", SMM would not display the metrics for the resource as the topic name contains the keyword "topic". If there exists another topic named as world whose name is a suffix of the first one, then SMM throws "For Input String" error in the UI.

### **Cannot read property 'indexOf' thrown when ETELatencyMetrics is disabled**

Bug ID: CSP-778

Apache JIRA: N/A

Component: Streams Messaging Manager

Summary: If latency metrics is disabled and you click the latency tab on the Topic Profile page in the SMM UI, then the following error appears: Cannot read property 'indexOf' of null.

### **SMM logs are not moving to the custom location**

Bug ID: CSP-838

Apache JIRA: N/A

Component: Streams Messaging Manager

Summary: SMM logs are not created in the custom location.

**SMM throws error if Service Monitor is not on CM host**

Bug ID: CSP-798

Apache JIRA: N/A

Component: Streams Messaging Manager

Summary: If the Cloudera Manager Service Monitor and Cloudera Manager Server are deployed on different hosts, SMM is unable to fetch metrics correctly. As a result, historic data for consumer offsets and lag are not displayed, only the latest data is available.

## Fixed issues in YARN

This section lists the issues that have been fixed since the previous version.

**COMPX-1403 Be able to view logs from YARN web UI v2**

Prior to this fix, YARN Application logs were accessible only from YARN WebUI1. With this fix, you can now access YARN logs in the YARN WebUI2.

**CDPD-10091: Upgrade jackson-databind to version 2.9.10.4 to resolve multiple CVEs**

Upgraded Jackson Databind to 2.9.10.4 to address the following CVEs: CVE-2020-9547, CVE-2020-11111, CVE-2020-10672, CVE-2020-10969, CVE-2020-11112, CVE-2020-9548, CVE-2020-9546, CVE-2020-10968, CVE-2020-10673, CVE-2020-11113.

**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.

## Fixed issues in Zeppelin

This section lists the issues that have been fixed since the previous version.

**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**

Workaround: 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.

Workaround: None.

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

Zeppelin 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.

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

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

## Fixed issues in ZooKeeper

This section lists the issues that have been fixed since the previous version.

**CDPD-10091: Upgrade jackson-databind to version 2.9.10.4 to resolve multiple CVEs**

Upgraded Jackson Databind to 2.9.10.4 to address the following CVEs: CVE-2020-9547, CVE-2020-11111, CVE-2020-10672, CVE-2020-10969, CVE-2020-11112, CVE-2020-9548, CVE-2020-9546, CVE-2020-10968, CVE-2020-10673, CVE-2020-11113.

#### **CDPD-10532: Update Log4j to address CVE-2019-17571**

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

#### **CDPD-7723: ZooKeeper - Upgrade to Jetty 9.4.26 to avoid CVEs**

ZooKeeper 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.

## Known Issues In Cloudera Runtime 7.1.1

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



**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

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

#### **Atlas notifications to Ranger are missing propagated classifications**

When an entity was updated or created, Atlas correctly propagates classifications from the parent table or tables to the new entity. However, when Atlas notifies Ranger of the new table, the notification does not include the propagated classification. If Ranger includes a tag-based access policy that corresponds to the Atlas classification, the policy will not be 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 propagates the PII classification from the parent table to the new table. The data Atlas sends to Ranger does not have the propagated "PII" classification, and therefore Ranger does not apply the tag-based access policy to the table.

Workaround: None.

Apache JIRA: Atlas-3806

#### **Incorrect attribute values in bulk import**

When importing Business Metadata attribute assignments, Atlas used only the last assigned attribute value instead of individual values for each entity in the import list. For example, setting Business Metadata attributes on entities as shown results in all entities to have the last value for the attributes: Processing.owner="FIN-admin" and Processing.track="standard".

```
EntityType,EntityUniqueAttributeValue,BusinessAttributeName,BusinessAttributeValue,EntityUniqueAttributeName[optional]
Table,customer_dim@cll,Processing.owner,"IT-admin"
Table,customer_dim@cll,Processing.track,"PII"
Table,log_fact_daily_mv@cll,Processing.owner,"IT-admin"
Table,log_fact_daily_mv@cll,Processing.track,"daily"
Table,time_dim@cll,Processing.owner,"FIN-admin"
Table,time_dim@cll,Processing.track,"standard"
```

Workaround: Include only one instance of a given attribute in a given import file.

Cloudera JIRA: CDPD-13199

#### **Migration progress bar not refreshed**

During the import stage of Cloudera Navigator to Apache Atlas migration, the migration progress bar does not correctly refresh the migration status. The Statistics page in the Atlas UI displays the correct details of the migration.

Workaround: None.

Cloudera JIRA: CDPD-12620

### **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.

Workaround: None.

Cloudera JIRA: CDPD-11790

### **Deleted Business Metadata attributes appear in Search Suggestions**

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

Workaround: None.

Cloudera JIRA: CDPD-10576

### **Suggestion order doesn't match search weights**

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

Workaround: None.

Cloudera JIRA: CDPD-10574

### **Hive Default Database Location Incorrect in Atlas Metadata**

The location of the default Hive database as reported through the HMS-Atlas plugin does not match the actual location of the database. This problem does not affect non-default databases.

Workaround: None.

Cloudera JIRA: CDPD-6042

### **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.

Workaround: None.

Cloudera JIRA: CDPD-5933, CDPD-5931

### **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.

Workaround: None.

Cloudera JIRA: CDPD-4762

### **Searches for Qualified Names with "@" doesn't fetch the correct results**

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

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

Cloudera JIRA: CDPD-4545

### **Missing Impala and Spark lineage between tables and their data files**

Atlas does not create lineage between Hive tables and their backing HDFS files for CTAS processes run in Impala or Spark.

Workaround: None.

Cloudera JIRA: CDP-5027, CDPD-3700, IMPALA-9070

### **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.

Workaround: None.

Cloudera JIRA: CDPD-3208

### **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.

Workaround: None.

Cloudera JIRA: CDPD-3160

### **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.

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

Cloudera JIRA: CDPD-3125

### **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.

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

Cloudera JIRA: CDPD-1892

### **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.

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

Workaround: None.

Cloudera JIRA: CDPD-1884

#### **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.

Workaround: None.

Cloudera JIRA: CDPD-1823

#### **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.

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

Cloudera JIRA: CDPD-1664

#### **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.

Workaround: None.

Cloudera JIRA: CDPD-922

#### **All Spark Queries from the Same Spark Session are Included in a Single Atlas Process**

A Spark session can include multiple queries. When Atlas reports the Spark metadata, it creates a single process entity to correspond to the Spark session. The result is that an Atlas lineage picture may show multiple input entities or multiple output entities from a process that are only related by the fact that they were included in operations in the same Spark session. The consequence of this behavior is that classifications will be propagated from any input entity to all output entities, even if the output entities aren't derived from the input entity.

Workaround: you can manually stop classification propagation to the inappropriate entity or choose not to propagate classifications that might be used in these scenarios.

Cloudera JIRA: CDPD-372

## **Known issues in Cruise Control**

This topic describes known issues for using Cruise Control in this release of Cloudera Runtime.

### 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 cannot collect metrics when topics are not used

Cruise Control will not collect metrics if the topics are not used by any consumer or producer, or if SRM is not active. In this case, the following error message is displayed:

```
10:02:53.046 AMWARN BrokerLoad
Broker 335 is missing 3/59 topics metrics and 14/133 leader partition metrics. Missing leader topics:
[mm2-offsets_CDFClusterPhoenix_internal, mm2-status_CDFClusterPhoenix_internal, srm-metrics_CDFClusterPhoenix_internal].
10:02:53.050 AMWARN SamplingUtils
Skip generating metric sample for broker 335 because there are not enough topic metrics to generate broker metrics.
```

Workaround:None

Cruise Control will generate proposals automatically as soon as topics are used or SRM becomes active.

### OPSAPS-58700: 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.

Workaround: 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.

### CDPD-10505: 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.

## Known Issues in DAS

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

- 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:

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

```
update das.report_scheduler_run_audit set status = 'FAILED' where status = 'READING';
```

- 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

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

### 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

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

**HBASE-24885: If an operator uses HBACK2 to invoke multiple `assigns` operations against one Region or happens to invoke HBACK2 `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**

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

#### HDFS encryption with HBase

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

Workaround: N/A

#### AccessController postOperation problems in asynchronous operations

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

- If a Delete Table fails for a reason other than missing permissions, the access rights are removed but the table may still exist and may be used again.
- If `hbaseAdmin.modifyTable()` is used to delete column families, the rights are not removed from the Access Control List (ACL) table. The `portOperation` is implemented only for `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.

Workaround: N/A

Apache Issue: [HBASE-6992](#)

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

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

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

Apache Issue: N/A

### 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](#)

**TSB 2022-569: HBase normalizer can cause table inconsistencies by merging non-adjacent regions**

The normalizer in HBase is a background job responsible for splitting or merging HBase regions to optimize the number of regions and the distribution of the size of the regions in HBase tables. Due to the bug described in HBASE-24376, the normalizer can cause region inconsistencies (region overlaps/holes) by merging non-adjacent regions.

**Upstream JIRA**

[HBASE-24376](#)

**Knowledge article**

For the latest update on this issue, see the corresponding Knowledge article: [TSB 2022-569: HBase normalizer can cause table inconsistencies by merging non-adjacent regions](#)

## Known Issues in HDFS

This topic describes known issues and unsupported features for using HDFS in this release of Cloudera Runtime. **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 2021-406: CVE-2020-9492 Hadoop filesystem bindings (ie: webhdfs) allows credential stealing**

WebHDFS clients might send SPNEGO authorization header to remote URL without proper verification. A maliciously crafted request can trigger services to send server credentials to a webhdfs path (ie: webhdfs://...) for capturing the service principal.

**Knowledge article**

For the latest update on this issue see the corresponding Knowledge article: [TSB-2021 406: CVE-2020-9492 Hadoop filesystem bindings \(ie: webhdfs\) allows credential stealing](#)

**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

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

**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.

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 "TBLS" ("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 "TBLS".
- 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-12301: Spark Hive Streaming using HWC fails**

Spark Hive Streaming using HWC can fail throwing a java.lang.NoSuchMethodException message.

Workaround: If you encounter this problem and need to use this capability, contact Cloudera for a fix.

**HiveServer Web UI displays incorrect data**

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

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

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

Workaround:

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. For example:

```
Name: tez.aux.uris
Value: file:///etc/hive/conf/ssl-client.xml
```

## 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

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

### 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 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.

### 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.

Workaround:



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.

Workaround:

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.

### **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.

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

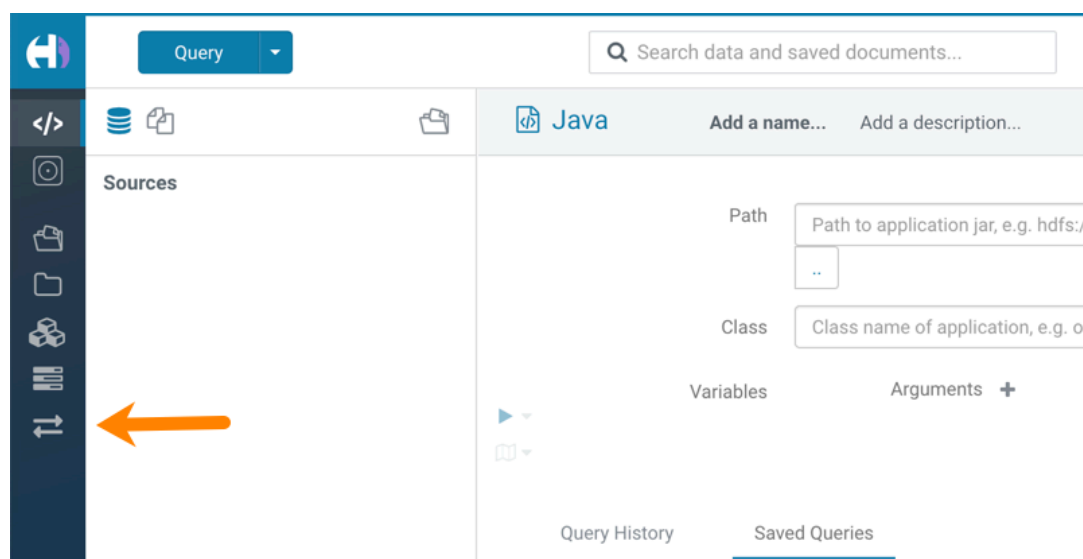
### **Hive and Impala query editors do not work with TLS 1.2**

Problem: If Hive or Impala engines are using TLS version 1.2 on your CDP cluster, then you won't be able to run queries from the Hue Hive or Impala query editor.

Workaround: You must apply the following patch: HUE-9508. Contact Cloudera Support for help on applying the software patch.

### **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 features

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

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

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



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

#### 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

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

### Impala known limitation when querying compacted tables

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

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

Use the [REFRESH/INVALIDATE](#) statements on the affected table to overcome the 'File does not exist' exception.

### 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.

Workaround: Restart the impalad.

Apache JIRA: HADOOP-15720

### Impala should tolerate bad locale settings

If the LC\_\* environment variables specify an unsupported locale, Impala does not start.

Workaround: Add LC\_ALL="C" to the environment settings for both the Impala daemon and the Statestore daemon.

Apache JIRA: IMPALA-532

### 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> >'
```

Workaround: 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
```

Apache JIRA: IMPALA-5605

#### 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".

Workaround: 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.

Apache JIRA: IMPALA-635

#### 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.

Workaround: 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.

Apache JIRA: IMPALA-691

#### Ranger audit logs for applying column masking policies missing

Impala is not producing these logs.

Workaround: None.

Apache JIRA: IMPALA-9350

#### 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.

Workaround: Remove trailing semicolon from the Avro schema.

Apache JIRA: IMPALA-1024

#### 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.

Workaround: Use the `RPAD()` function to blank-pad literals compared with CHAR columns to the expected length.



Apache JIRA: IMPALA-3094

### **Breakpad minidumps can be very large when the thread count is high**

The size of the breakpad minidump files grows linearly with the number of threads. By default, each thread adds 8 KB to the minidump size. Minidump files could consume significant disk space when the daemons have a high number of threads.

Workaround: 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.

Apache JIRA: IMPALA-3509

### **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.

Workaround: 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.

Apache JIRA: IMPALA-4978

### **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.

Workaround:

Apache JIRA: IMPALA-6671

### **Impala does not support Heimdal Kerberos**

Apache JIRA: IMPALA-7072

### **CDPD-28139: Set spark.hadoop.hive.stats.autogather to false by default**

As an Impala user, if you submit a query against a table containing data ingested using Spark and you are concerned about the quality of the query plan, you must run `COMPUTE STATS` against such a table in any case after an ETL operation because `numRows` created by Spark could be incorrect. Also, use other stats computed by `COMPUTE STATS`, e.g., Number of Distinct Values (NDV) and NULL count for good selectivity estimates.

For example, when a user ingests data from a file into a partition of an existing table using Spark, if `spark.hadoop.hive.stats.autogather` is not set to false explicitly, `numRows` associated with this partition would be 0 even though there is at least one row in the file. To avoid this, the workaround is to set `"spark.hadoop.hive.stats.autogather=false"` in the "Spark Client Advanced Configuration Snippet (Safety Valve) for spark-conf/spark-defaults.conf" in Spark's CM Configuration section.

## **Technical Service Bulletins**

### **TSB-2021-485: 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](#)

#### **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 compactations 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 compactations 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

This topic describes known issues, unsupported features and limitations for using Kafka in this release of Cloudera Runtime.

### Known Issues

#### OPSAPS-59553: SMM's bootstrap server config should be updated based on Kafka's listeners

SMM does not show any metrics for Kafka or Kafka Connect when multiple listeners are set in Kafka.



Workaround: SMM cannot identify multiple listeners and still points to bootstrap server using the default broker port (9093 for SASL\_SSL). You would have to override bootstrap server URL (hostname:port as set in the listeners for broker) in the following path:

Cloudera Manager > SMM > Configuration > Streams Messaging Manager Rest Admin Server Advanced Configuration Snippet (Safety Valve) for streams-messaging-manager.yaml > Save Changes > Restart SMM.

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

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

Workaround: 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

Workaround: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.

Workaround: 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.

Workaround: 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.

Workaround: 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.

Workaround: Configure brokers and clients with ssl.secure.random.implementation = SHA1PRNG. It often reduces this degradation drastically, but its effect is CPU and JVM dependent.

Apache JIRA: KAFKA-2561

### **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.

Workaround: None.

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

If the Kafka Broker Advanced Configuration Snippet (Safety Valve) for ssl.properties property contains configuration that has dollar signs, the configuration is not propagated to Kafka brokers correctly.

Workaround: 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.

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

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

If a large number of topic partitions are created on a cluster, the Cloudera Manager Agent can generate a high CPU load. This is caused by the Kafka metric collector adapter carrying out excessive regex matching.

Workaround: None.

**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.

Workaround: 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 Port Streams Messaging Manager property. The port configured in this property must match the port configured for the Kafka Connect role.

**Kafka Connect fails to communicate with secured Schema Registry**

Kafka Connect connectors establish a connection with a Schema Registry server if they are configured to use the AvroConverter. If the Schema Registry server is Kerberos enabled, a valid JAAS configuration is required to establish a connection with the server. The JAAS configuration is specified with the \*.converter.sasl.jaas.config connector property. However, due to an underlying issue in Schema Registry, this is not possible. As a result, connectors might fail to connect to Kerberos enabled Schema Registry servers.

Workaround: Manually create a JAAS configuration that includes a valid RegistryClient entry and add it to the KAFKA\_OPTS environment variable.



**Important:** Complete these steps for each Kafka Connect host.

1. Create a copy of the Kafka Connect principal keytab and deploy it on the host. For example:

```
/etc/kafka/conf/kafka-connect.keytab
```

2. Create a JAAS configuration file containing a RegistryClient entry. For example:

```
RegistryClient {
  com.sun.security.auth.module.Krb5LoginModule required
  doNotPrompt=true
  useKeyTab=true
  useTicketCache=false
  storeKey=true
  keyTab="/etc/kafka/conf/kafka-connect.keytab"
  principal="kafka/host1.cloudera.example.com@CLOUDERA.EXAMPLE.COM";
};
```

Ensure that you replace the value of keyTab with the full path to the copy of the keytab file you created in Step 1.

3. Ensure that both the keytab and the JAAS file are accessible by the Kafka Connect process user.
4. In Cloudera Manager, go to KafkaConfiguration and find the Kafka Connect Environment Advanced Configuration Snippet (Safety Valve) property.
5. Add the following property to the advanced configuration snippet.

```
Key: KAFKA_OPTS
Value: -Djava.security.auth.login.config=/etc/kafka/conf/kafka-connect-jaas.conf
```

Ensure that you replace the value of -Djava.security.auth.login.config with the full path to the JAAS file you created in Step 2. In this example, the JAAS file is called kafka-connect-jaas.conf and is located in /etc/kafka/conf/.

6. Click Save Changes.
7. Restart the Kafka Connect roles.

## Unsupported Features

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

- Only Java based clients are supported. Clients developed with C, C++, Python, .NET 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

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

**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.

Workaround: To prevent additional access to Atlas, close all browser windows and exit the browser.

**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

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

- 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

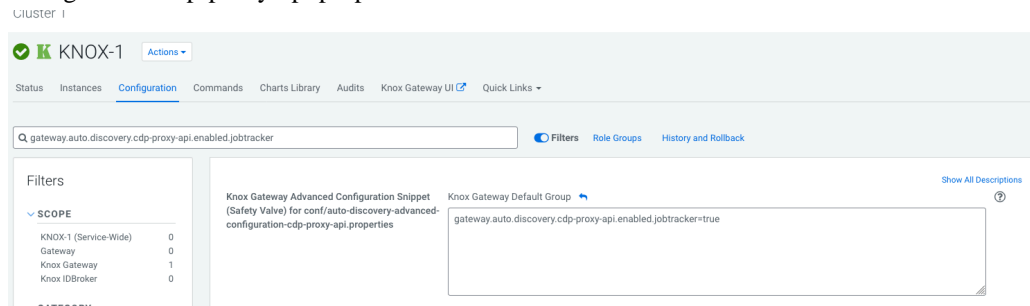
This topic describes known issues and unsupported features for using Oozie in this release of Cloudera Runtime.

### Oozie's limited Knox support

By default, Oozie does not work on a Knox enabled cluster as Knox is unable to replace the resource-manager property for Oozie.

Workaround: Oozie 5 contains a backward-compatibility implementation where users can continue to use the job-tracker property for specifying the resource-manager's address. This mode is enabled by default in Oozie and when Knox is installed, you must manually configure an additional property in Cloudera Manager to enable the same functionality in Knox.

1. In Cloudera Manager, select the Knox service.
2. Click the Configuration tab.
3. Search for auto-discovery-advanced-configuration-cdp-proxy-api.
4. Set gateway.auto.discovery.cdp-proxy-api.enabled.jobtracker=true as the value in the Knox Gateway Advanced Configuration Snippet (Safety Valve) for conf/auto-discovery-advanced-configuration-cdp-proxy-api.properties field.



Wait until Cloudera Manager recognizes the stale configuration.

5. Click the newly appeared Refresh button next to the Knox service and follow the on-screen instructions to restart/refresh Knox.

With this workaround, a limited Knox usage can be achieved in Oozie where the resource-manager property still does not work on a Knox enabled cluster, but you can achieve the same behaviour through the job-tracker property.

### 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.

Workaround: When the JobHistory server is running again, use the resume command to inform Oozie to continue the workflow from the point at which it left off.

### Unsupported Feature

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

- Non-support for Pig action (CDPD-1070)
- Conditional coordinator input logic

Cloudera does not support using Derby database with Oozie. You can use it for testing or debugging purposes, but Cloudera does not recommend using it in production environments. This could cause failures while upgrading from CDH to CDP.

### Technical Service Bulletins

#### TSB 2021-467: Race condition in Apache Oozie Sharelib upload

There is a race condition in Apache Oozie OozieSharelibCLI which allows a malicious attacker to replace the files in Oozie's sharelib during its creation. A race condition in OozieSharelibCLI allows an attacker to replace the contents of the sharelib.

#### Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-467: Race condition in Apache Oozie Sharelib upload](#)

## Ozone

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

#### A user with ALL Access in Ranger cannot list volumes created by other users

When `ozone.acl.enabled=True` and `ozone.administrators` are defined, ACL checks such as volume create and list volume are not sent to the configured authorizer plug-in such as Ranger or NativeOzoneAuthorizer; instead, they are based on the Ozone Manager's `ozone.administrators` configuration.

As a result, if you set the authorizer policy to allow certain user to create or list volumes, the request is not honored.

Workaround: Ensure that admin operations such as create volume and list all volumes are allowed only for those users defined in the `ozone.administrators` configuration.

Cloudera JIRA: CDPD-12096

#### The Recon web user interface shows incomplete information about Ozone volumes, buckets, and keys

In a secure cluster with High Availability for Ozone Manager enabled, if Recon is not configured with the correct server principal of the Ozone Manager, it cannot receive updates from Ozone Manager on a regular basis. Therefore, the Recon web user interface shows incomplete information about volumes, buckets, and keys.

Workaround: Add the following Recon configuration property using the Ozone Recon Advanced Configuration Snippet (Safety Valve) for `ozone-conf/ozone-site.xml` configuration parameter from Cloudera Manager: `ozone.om.kerberos.principal=<comma separated list of all the Ozone Manager principals configured for the cluster>`.

## Known Issues in Apache Ranger

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

#### 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.

Workaround:

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.

Workaround:

Use only unique key names when creating keys.

## Known Issues in Apache Ranger KMS

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

#### **Documentation for Ranger KMS is not complete**

Ranger KMS has been added to this release, but the applicable topics under "Encrypting Data at Rest" and "HDFS Transparent Data Encryption" are still being updated to reflect the replacement of Key Trustee KMS with Ranger KMS.

Workaround:

None, but these topics will be completed as soon as possible.

## Known Issues in Schema Registry

This topic describes known issues, unsupported features and limitations for using Schema Registry in this release of Cloudera Runtime.

#### **You cannot upload a JAR file through the Schema Registry UI if you use Firefox or Internet Explorer**

If you use Firefox or Internet Explorer to access the Schema Registry UI, you will get an Invalid File Type error when you try to upload a JAR file from the SchemaRegistry UI.

Workaround: Use the Chrome browser to upload a JAR file. Alternatively, you can upload a JAR file using the REST API.

## Known Issues in Cloudera Search

This topic describes known issues and unsupported features for using Cloudera Search in this release of Cloudera Runtime.

#### **Known Issues**

##### **Cloudera Bug ID:**

CDPD-20577

##### **Summary:**

Splitshard of HDFS index checks local filesystem and fails

**Description:**

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.

**Workaround:**

None

**Cloudera Bug ID:**

OPSAPS-58059

**Summary:**

Solr log rotation counts the number of retained log files daily instead of globally

**Description:**

With CDP 7.1.1, Search moved to Log4Jv2. This has affected Solr log rotation behavior in an unwanted way. With the default configuration, Solr log file names include a date and a running index, for example: solr-cmf-solr-SOLR\_SERVER-solrserver-1.my.corporation.com.log.out.2020-08-31-9. The number of retained log files is configured in Cloudera Manager, however the configured number now applies for each day, instead of applying globally for all log files of the particular server.

**Workaround:**

Using Cloudera Manager, edit the Solr Server Logging Advanced Configuration Snippet (Safety Valve) property of your Solr service and add a new line containing: appender.DRFA.filePattern=\${log.dir}/\${log.file}.%i

**Cloudera Bug ID:**

DOCS-5717

**Summary:**

Lucene index handling limitation

**Description:**

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

**Workaround:**

There is no workaround, you need to reindex collections.

**Cloudera Bug ID:**

CDH-82042

**Summary:**

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

**Description:**

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.

**Workaround:**



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

**Cloudera Bug ID:**

CDH-66345

**Summary:**

Solr SQL, Graph, and Stream Handlers are Disabled if Collection Uses Document-Level Security

**Description:**

The Solr SQL, Graph, and Stream handlers do not support document-level security, and are disabled if document-level security is enabled on the collection. If necessary, these handlers can be re-enabled by setting the following Java system properties, but document-level security is not enforced for these handlers:

- SQL: solr.sentry.enableSqlQuery=true
- Graph: solr.sentry.enableGraphQuery=true
- Stream: solr.sentry.enableStreams=true

**Workaround:**

None.

**Cloudera Bug ID:**

CDH-34050

**Summary:**

Collection Creation No Longer Supports Automatically Selecting A Configuration If Only One Exists

**Description:**

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.

**Workaround:**

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

**Cloudera Bug ID:**

CDH-22190

**Summary:**

CrunchIndexerTool which includes Spark indexer requires specific input file format specifications

**Description:**

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.

**Workaround:**

None.

**Cloudera Bug ID:**

CDH-19923

**Summary:**

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

**Description:**

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.

**Workaround:**

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.

**Cloudera Bug ID:**

CDH-26856

**Summary:**

Field value class guessing and Automatic schema field addition are not supported with the MapReduceIndexerTool nor with the HBaseMapReduceIndexerTool

**Description:**

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.

**Workaround:**

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.

**Cloudera Bug ID:**

CDH-19407

**Summary:**

The Browse and Spell Request Handlers are not enabled in schemaless mode

**Description:**

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.

**Workaround:**

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.

**Cloudera Bug ID:**

CDH-17978

**Summary:**

Enabling blockcache writing may result in unusable indexes

**Description:**

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.

**Workaround:**

None.

**Cloudera Bug ID:**

CDH-58276

**Summary:**

Users with insufficient Solr permissions may receive a "Page Loading" message from the Solr Web Admin UI

**Description:**

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.

**Workaround:**

None.

**Cloudera Bug ID:**

CDH-15441

**Summary:**

Using `MapReduceIndexerTool` or `HBaseMapReduceIndexerTool` multiple times may produce duplicate entries in a collection

**Description:**

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.

**Workaround:**

To avoid this issue, use `HBaseMapReduceIndexerTool` with zero reducers. This must be done without Kerberos.

**Cloudera Bug ID:**

CDH-58694

**Summary:**

Deleting collections might fail if hosts are unavailable

**Description:**

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.

**Workaround:**

Ensure all hosts are online before deleting collections.

**Cloudera Bug ID:**

CDH-58694

**Summary:**

Saving search results is not supported

**Description:**

Cloudera Search does not support the ability to save search results.

Workaround:

None.

**Cloudera Bug ID:**

CDH-11357

**Summary:**

HDFS Federation is not supported

**Description:**

Cloudera Search does not support HDFS Federation.

**Workaround:**

None.

**Cloudera Bug ID:**

CDPD-4139

**Summary:**

Collection state goes down after Solr SSL

**Description:**

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.

**Workaround:**

[Recreate the collection after enabling TLS.](#)

**Cloudera Bug ID:**

CDPD-13923

**Summary:**

Every Configset is Untrusted Without Kerberos

**Description:**

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.

**Workaround:**

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

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

**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 Spark HBase Connector (SHC) from HDP and the hbase-spark module from CDH are not supported.

Workaround: Migrate to the Apache HBase Connectors integration for Apache Spark (hbase-connectors/spark) available in CDP. More details on the integration for working with HBase data from Spark in CDP is available in the Cloudera Community article, [HBase and Spark in CDP](#).

#### **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
```

Workaround: These errors can be safely ignored.

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

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

Workaround: Create the view using Hive or Impala.

#### **CDPD-11720: HDFS ACLs not set on Hive external warehouse if Impala is not on cluster**

If Impala is installed on the cluster, Impala sets HDFS ACLs on both the managed and external Hive warehouse. This allows Spark to write to tables created in the Hive external warehouse. If Impala is not installed, then these HDFS ACLs are not set, and Spark is not able to write to external tables created by Hive.

Workaround: Set HDFS ACLs manually.

#### **CDPD-12622: Sentry GRANTS given during pre-migration to a role does not work post-migration in Spark Shell**

Spark requires SELECT privileges on the default database, regardless of the databases referenced in the query.

Workaround: Add SELECT privileges in Ranger to the default database for all users who will run Spark queries.

### **Technical Service Bulletins**

#### **TSB 2021-441: CDP Powered by Apache Spark may incorrectly read/write pre-Gregorian timestamps**

Spark may incorrectly read or write TIMESTAMP data for values before the start of the Gregorian calendar ('1582-10-15 00:00:00.0'). This could happen when Spark is:

- Using dynamic partition inserts
- Reading or writing from an ORC table when the:
  - spark.sql.hive.convertMetastoreOrc property is set to false. Its default value is true.
  - spark.sql.hive.convertMetastoreOrc property is set to true but the spark.sql.orc.impl property is set to hive. Its default is native.
- Reading or writing from a Parquet table when the:
  - spark.sql.hive.convertMetastoreParquet property is set to false. Its default value is true.

#### **Knowledge article**

For the latest update on this issue see the corresponding Knowledge article: [TSB 2021-441: Spark may incorrectly read/write pre-Gregorian timestamps](#)

## Known Issues in Streams Replication Manager

This topic describes known issues for using Streams Replication Manager in this release of Cloudera Runtime.

### Known Issues

#### 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.

N/A.

#### 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 whitelist 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.

N/A.

#### 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 flow 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.

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

When polling the srm-control.<alias>.internal internal configuration topic, it may happen that not all records are returned at once. It can happen that the first poll only returns a single message. Remaining messages are only returned on a subsequent poll. As a result, only parts of the configuration are picked up. This causes the srm-control tool and the SRM driver to behave erratically as they are unable to read the full white and blacklists from the configuration topic.

N/A.

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

If a topic was originally replicated but was later blacklisted, it will still appear as a replicated topic under the /remote-topics REST API endpoint. As a result, if a call is made to this endpoint, the blacklisted topic will be included in the response. Additionally, the blacklisted 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.



N/A.

**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.

N/A.

**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 in Streams Messaging Manager

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

**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 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**

Problem: Importing an Avro file into S3 with HCat fails with Delegation Token not available.

CDPD-3089

### **Parquet columns inadvertently renamed**

Problem: Column names that start with a number are renamed when you use the --as-parquetfile option to import data.

Workaround: Prepend column names in Parquet tables with one or more letters or underscore characters.

Apache JIRA: None

### **Importing Parquet files might cause out-of-memory (OOM) errors**

Problem: 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.

PARQUET-99

## Known Issues in MapReduce and YARN

This topic describes known issues, unsupported features and limitations for using MapReduce and YARN in this release of Cloudera Runtime.

### Known Issues

**OPSAPS-56577: If a Kerberos principal other than "yarn" is configured for the YARN service, then Cloudera Manager will erroneously skip adding the custom principal to the YARN keytab, causing YARN to fail to start due to a Kerberos authentication failure. This also affects Ambari to Cloudera Manager migrations, if Ambari is configured with a principal other than "yarn" for the YARN service. A similar issue affects Hive, when using Hive LLAP.**

Workaround: You must specify "yarn" as the Kerberos principal for YARN, and specify "hive" as the Kerberos principal for Hive. When performing an Ambari to Cloudera Manager migration, set the principals for both services to those values before performing the migration.

### Fair Scheduler to Capacity Scheduler migration - fs2cs tool

fs2cs tool does not convert all Fair Scheduler queue configurations to Capacity Scheduler queue configurations.

Workaround: You must manually configure the queue configurations which are not converted by the fs2cs tool. For information about using the fs2cs tool and its limitations, see [Fair Scheduler to Capacity Scheduler transition](#).

### CDPD-12123: HDFS replication performance is either slowed down or not on the expected lines with CDH/CM 7.1.1.

Post-upgrade, queue which the user is running workloads cannot grow beyond the configured capacity till its maximum capacity.

Workaround: Once the cluster is upgraded to CDP Private Cloud Base 7.1.1, user may need to tune `yarn.scheduler.capacity.<queuepath>.user-limit-factor` to a value greater than 1. This configuration enables the queue usage to grow beyond its configured capacity, till its maximum capacity configured.

### DOCS-5966: Third party applications do not launch if MapReduce framework path is not included in the client configuration

MapReduce application framework is loaded from HDFS instead of being present on the NodeManagers. By default, the `mapreduce.application.framework.path` property is set to the appropriate value, but third party applications with their own configurations will not launch.

Workaround: Set the `mapreduce.application.framework.path` property to the appropriate configuration for third party applications.

### JobHistory URL mismatch after server relocation

After moving the JobHistory Server to a new host, the URLs listed for the JobHistory Server on the ResourceManager web UI still point to the old JobHistory Server. This affects existing jobs only. New jobs started after the move are not affected.

Workaround: For any existing jobs that have the incorrect JobHistory Server URL, there is no option other than to allow the jobs to roll off the history over time. For new jobs, make sure that all clients have the updated `mapred-site.xml` that references the correct JobHistory Server.

### CDH-49165: History link in ResourceManager web UI broken for killed Spark applications

When a Spark application is killed, the history link in the ResourceManager web UI does not work.

Workaround: 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.

Workaround: 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.

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

**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 operations are failing. Queue Manager says 0 queues are configured and several failures are present. That is because ZooKeeper configuration store is not enabled.

Workaround:

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**

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.

Workaround: Reload the Queue Manager page manually.

**COMPX-3134: Yarn applications can get stuck due to a NullPointerException in Capacity Scheduler**

If you enable Asynchronous scheduling (`yarn.scheduler.capacity.schedule-asynchronously.enable=true`) in capacity scheduler, there is an edge-case where NullPointerException can cause the scheduler thread to exit and the applications get stuck without allocated resources. This can be recognized by NullPointerException thrown by the capacity scheduler.

Workaround: Restart the ResourceManager and check if the resources are allocated to the applications that were stuck.

**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 Enabl 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:

- GPU support for Docker
- Hadoop Pipes
- Fair Scheduler
- Application Timeline Server (ATS 2 and ATS 1.5)
- Container Resizing
- Distributed or Centralized Allocation of Opportunistic Containers
- Distributed Scheduling
- Native Services
- Pluggable Scheduler Configuration
- Queue Priority Support
- Reservation REST APIs
- Resource Estimator Service
- Resource Profiles
- (non-Zookeeper) ResourceManager State Store
- Shared Cache
- YARN Federation
- Rolling Log Aggregation
- Docker on YARN (DockerContainerExecutor) on Data Hub clusters
- Moving jobs between queues
- Dynamic Resource Pools

**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**

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

**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.

Workaround: 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.

Workaround: Close the browser.

**Known Issues in Apache ZooKeeper**

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

**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.

Workaround:

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.zookeeper.ClientCnxnSocketNetty -Dzookeeper.ssl.keyStore.location=<path to your configured keystore> -Dzookeeper.ssl.keyStore.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.secure=true" zookeeper-client -server <your.zookeeper.server-1>:2182,<your.zookeeper.server-2>:2182,<your.zookeeper.server-3>:2182
```

**ZooKeeper cluster can be slow to start if QuorumSSL is enabled without QuorumSASL**

QuorumSSL (Secure ZooKeeper) is enabled by default if AutoTLS is enabled. If QuorumSSL is enabled without QuorumSASL (Server to server SASL authentication), then the ZooKeeper cluster can be slow to start due to some known ZooKeeper limitations.

Workaround:

Ensure, that QuorumSSL is enabled only if QuorumSASL is also enabled:

1. In Cloudera Manager, select the ZooKeeper service.
2. Click the Configuration tab.
3. Search for SSL.
4. Find and use the Enable TLS/SSL for ZooKeeper property to enable QuorumSSL.
5. Search for SASL.
6. Find and use the Enable Server to Server SASL Authentication property to enable QuorumSASL.
7. Click Save Changes.