

Release Notes

Date published:

Date modified:

CLOUDERA

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

Overview.....	5
Using the Cloudera Runtime Maven Repository.....	5
Maven Artifacts for Cloudera Runtime 7.0.3.0.....	5
What's New In Cloudera Runtime 7.0.3.....	23
What's New in Apache Atlas.....	23
What's new in DAS.....	24
What's New in Apache HBase.....	24
What's New in Apache Hadoop HDFS.....	24
What's New in Apache Hive.....	24
What's New in Hue.....	24
What's New in Apache Impala.....	26
What's New in Apache Kafka.....	30
What's New in Apache Kudu.....	30
What's New in Apache Oozie.....	32
What's New in Apache Hadoop Ozone.....	32
What's New in Apache Phoenix.....	32
What's New in Apache Ranger.....	32
What's New in Cloudera Search.....	33
What's New in Apache Spark.....	33
What's New in Apache Hadoop YARN.....	33
What's New in Apache ZooKeeper.....	33
Known Issues In Cloudera Runtime 7.0.3.....	34
Known Issues in Apache Atlas.....	34
Known Issues in Apache Accumulo.....	37
Known Issues in DAS.....	37
Known Issues in Apache HBase.....	38
Known Issues in HDFS.....	40
Known Issues in Apache Hive.....	40
Known Issues in Hue.....	42
Known Issues in Apache Impala.....	43
Known Issues in Apache Kafka.....	45
Known Issues in Apache Kudu.....	46
Known Issues in Apache Oozie.....	46
Known Issues in Cloudera Search.....	47
Known Issues in Apache Spark.....	50
Known Issues for Apache Sqoop.....	50
Known Issues in MapReduce and YARN.....	51
Cloudera Runtime Component Versions.....	54
Updated HDP Components.....	55

Updated CDH Components.....	55
------------------------------------	-----------

Overview

Cloudera Runtime release notes cover the new features, known issues, fixed issues, and the runtime component versions in this version of Cloudera Runtime.

Related Information

[What's New In Cloudera Runtime 7.0.3](#)

[Known Issues In Cloudera Runtime 7.0.3](#)

[Cloudera Runtime Component Versions](#)

[Updated HDP Components](#)

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.



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 7.0.3.0

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

Project	groupId	artifactId	version
Atlas	org.apache.atlas	atlas-authorization	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-classification-updater	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-client-common	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-client-v1	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-client-v2	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-common	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-distro	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-docs	2.0.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.atlas	atlas-graphdb-api	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-graphdb-common	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-graphdb-janus	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-intg	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-janusgraph-hbase2	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-notification	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-plugin-classloader	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-repository	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-server-api	2.0.0.7.0.3.0-79
	org.apache.atlas	atlas-testtools	2.0.0.7.0.3.0-79
	org.apache.atlas	hbase-bridge	2.0.0.7.0.3.0-79
	org.apache.atlas	hbase-bridge-shim	2.0.0.7.0.3.0-79
	org.apache.atlas	hbase-testing-util	2.0.0.7.0.3.0-79
	org.apache.atlas	hdfs-model	2.0.0.7.0.3.0-79
	org.apache.atlas	hive-bridge	2.0.0.7.0.3.0-79
	org.apache.atlas	hive-bridge-shim	2.0.0.7.0.3.0-79
	org.apache.atlas	impala-bridge	2.0.0.7.0.3.0-79
	org.apache.atlas	impala-bridge-shim	2.0.0.7.0.3.0-79
	org.apache.atlas	impala-hook-api	2.0.0.7.0.3.0-79
	org.apache.atlas	kafka-bridge	2.0.0.7.0.3.0-79
	org.apache.atlas	sqoop-bridge	2.0.0.7.0.3.0-79
	org.apache.atlas	sqoop-bridge-shim	2.0.0.7.0.3.0-79
	org.apache.atlas	storm-bridge	2.0.0.7.0.3.0-79
	org.apache.atlas	storm-bridge-shim	2.0.0.7.0.3.0-79
Apache Avro	org.apache.avro	avro	1.8.2.7.0.3.0-79
	org.apache.avro	avro-compiler	1.8.2.7.0.3.0-79
	org.apache.avro	avro-guava-dependencies	1.8.2.7.0.3.0-79
	org.apache.avro	avro-ipc	1.8.2.7.0.3.0-79
	org.apache.avro	avro-mapred	1.8.2.7.0.3.0-79
	org.apache.avro	avro-maven-plugin	1.8.2.7.0.3.0-79
	org.apache.avro	avro-protobuf	1.8.2.7.0.3.0-79
	org.apache.avro	avro-service-archetype	1.8.2.7.0.3.0-79
	org.apache.avro	avro-thrift	1.8.2.7.0.3.0-79
	org.apache.avro	avro-tools	1.8.2.7.0.3.0-79
	org.apache.avro	trevni-avro	1.8.2.7.0.3.0-79
	org.apache.avro	trevni-core	1.8.2.7.0.3.0-79
Calcite	org.apache.calcite	calcite-babel	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-cassandra	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-core	1.19.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.calcite	calcite-druid	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-elasticsearch	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-example-csv	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-example-function	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-file	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-geode	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-linq4j	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-mongodb	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-pig	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-piglet	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-plus	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-server	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-spark	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-splunk	1.19.0.7.0.3.0-79
	org.apache.calcite	calcite-ubenchmark	1.19.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica	1.10.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica-core	1.10.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica-metrics	1.10.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica-metrics-dropwizardmetrics3	1.10.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica-noop-driver	1.10.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica-server	1.10.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica-standalone-server	1.10.0.7.0.3.0-79
	org.apache.calcite.avatica	avatica-tck	1.10.0.7.0.3.0-79
Apache Crunch	org.apache.crunch	crunch-archetype	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-contrib	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-core	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-examples	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-hbase	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-hive	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-scrunch	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-spark	0.11.0.7.0.3.0-79
	org.apache.crunch	crunch-test	0.11.0.7.0.3.0-79
Druid	org.apache.druid	druid-aws-common	0.15.1.7.0.3.0-79
	org.apache.druid	druid-benchmarks	0.15.1.7.0.3.0-79
	org.apache.druid	druid-console	0.15.1.7.0.3.0-79
	org.apache.druid	druid-core	0.15.1.7.0.3.0-79
	org.apache.druid	druid-gcp-common	0.15.1.7.0.3.0-79
	org.apache.druid	druid-hll	0.15.1.7.0.3.0-79
	org.apache.druid	druid-indexing-hadoop	0.15.1.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.druid	druid-indexing-service	0.15.1.7.0.3.0-79
	org.apache.druid	druid-integration-tests	0.15.1.7.0.3.0-79
	org.apache.druid	druid-processing	0.15.1.7.0.3.0-79
	org.apache.druid	druid-server	0.15.1.7.0.3.0-79
	org.apache.druid	druid-services	0.15.1.7.0.3.0-79
	org.apache.druid	druid-sql	0.15.1.7.0.3.0-79
	org.apache.druid	extendedset	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hbase-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hadoop-basic-security	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hadoop-bloom-filter	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hatias-tasketches	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hc2-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hive-examples	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hogwe-googole-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-hdfs-storage	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-histogram	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-kafka-eight	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-kafka-extraction-namespac	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-kafka-indexing-service	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-kerberos	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-kinesis-indexing-service	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-lookups-cached-global	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-lookups-cached-single	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-morc-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-parquet-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-protobuf-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-s3-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-stats	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-sql-metadata-storage	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-sql-metadata-storage	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-ssl-client-sslcontext	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-ssl-contexts-emitter	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-ssl-contexts-emitter	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-szoozoo-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-szoozoo-storage	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-tkconfdfiles-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-tomcat-mount	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-trimble-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.extensions	druid-trimble-light-simple-consumer	0.15.1.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.druid.ex	druid-contrib-scheduler	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-average-query	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-emitter	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-hq	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-cache	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-hq	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-extensions	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-min-max	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-columns	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-view-maintenance	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-view-selection	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib-data-storage	0.15.1.7.0.3.0-79
	org.apache.druid.ex	druid-contrib	0.15.1.7.0.3.0-79
GCS Connector	com.google.cloud.bigtable	bigtable-gcs-connector	1.9.10.7.0.3.0-79
	com.google.cloud.bigtable	bigtable-connector	1.9.10.7.0.3.0-79
	com.google.cloud.bigtable	bigtable	1.9.10.7.0.3.0-79
	com.google.cloud.bigtable	bigtable	1.9.10.7.0.3.0-79
	com.google.cloud.bigtable	bigtable	1.9.10.7.0.3.0-79
Apache Hadoop	org.apache.hadoop	hadoop-aliyun	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-annotations	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-archive-logs	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-archives	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-assemblies	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-auth	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-aws	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-azure	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-azure-datalake	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-build-tools	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-client	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-client-api	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-client-integration-tests	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-client-minicluster	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-client-runtime	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-cloud-storage	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-common	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-datajoin	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-distcp	3.1.1.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-extras	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-fs2img	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-gridmix	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-client	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-common	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-config	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-container-service	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-docs	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-server-framework	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-server-scm	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdds-tools	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-hdfs	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-hdfs-client	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-hdfs-httpfs	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-hdfs-native-client	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-hdfs-nfs	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-hdfs-rbf	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-kafka	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-kms	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-app	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-common	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-core	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-hs	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-hs-plugins	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-jobclient	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-nativetask	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-shuffle	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-client-uploader	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-mapreduce-examples	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-maven-plugins	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-minicluster	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-minikdc	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-nfs	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-openstack	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-ozon-client	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozon-common	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozon-csi	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozon-datanode	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozon-filesystem	0.4.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.hadoop	hadoop-ozone-fsfilesystem-lib-current	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-fsfilesystem-lib-legacy	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-insight	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-integration-test	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-network-tests	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-ozone-manager	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-recon	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-reconcodegen	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-s3gateway	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-tools	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-ozone-upgrade	0.4.0.7.0.3.0-79
	org.apache.hadoop	hadoop-resourceestimator	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-rumen	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-sls	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-streaming	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-tools-dist	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-api	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-applications-distributedshell	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-applications-unmanaged-am-launcher	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-client	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-common	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-registry	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-applicationhistoryservice	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-common	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-nodemanager	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-resourcemanager	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-router	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-sharedcachemanager	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-tests	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-timeline-pluginstorage	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-timelineservice	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-client	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-common	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-server-2	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-tests	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-server-web-proxy	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-services-api	3.1.1.7.0.3.0-79
	org.apache.hadoop	hadoop-yarn-services-core	3.1.1.7.0.3.0-79
Apache HBase	org.apache.hbase	hbase-annotations	2.2.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.hbase	hbase-checkstyle	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-client	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-client-project	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-common	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-endpoint	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-examples	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-external-blockcache	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-hadoop-compat	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-hadoop2-compat	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-hbtop	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-http	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-it	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-mapreduce	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-metrics	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-metrics-api	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-procedure	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-protocol	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-protocol-shaded	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-replication	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-resource-bundle	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-rest	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-rsgroup	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-server	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-shaded-client	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-shaded-client-byo-hadoop	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-shaded-client-project	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-shaded-mapreduce	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-shaded-testing-util	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-shaded-testing-util-tester	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-shell	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-testing-util	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-thrift	2.2.0.7.0.3.0-79
	org.apache.hbase	hbase-zookeeper	2.2.0.7.0.3.0-79
	org.apache.hbase.contrib	hbase-kafka-model	1.0.0.7.0.3.0-79
	org.apache.hbase.contrib	hbase-kafka-proxy	1.0.0.7.0.3.0-79
	org.apache.hbase.contrib	hbase-spark	1.0.0.7.0.3.0-79
	org.apache.hbase.contrib	hbase-spark-it	1.0.0.7.0.3.0-79
	org.apache.hbase.contrib	hbase-spark-protocol	1.0.0.7.0.3.0-79
	org.apache.hbase.contrib	hbase-spark-protocol-shaded	1.0.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.hbase.filecache	hbase-cross	1.0.0.7.0.3.0-79
Apache Hive	org.apache.hive	hive-accumulo-handler	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-beeline	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-classification	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-cli	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-common	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-contrib	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-druid-handler	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-exec	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-hbase-handler	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-hplsql	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-jdbc	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-jdbc-handler	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-kryo-registrator	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-kudu-handler	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-llap-client	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-llap-common	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-llap-ext-client	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-llap-server	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-llap-tez	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-metastore	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-pre-upgrade	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-serde	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-service	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-service-rpc	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-shims	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-spark-client	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-standalone-metastore	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-storage-api	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-streaming	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-testutils	3.1.2000.7.0.3.0-79
	org.apache.hive	hive-vector-code-gen	3.1.2000.7.0.3.0-79
	org.apache.hive	kafka-handler	3.1.2000.7.0.3.0-79
	org.apache.hive.hcatalog	hive-hcatalog-core	3.1.2000.7.0.3.0-79
	org.apache.hive.hcatalog	hive-hcatalog-pig-adapter	3.1.2000.7.0.3.0-79
	org.apache.hive.hcatalog	hive-hcatalog-server-extensions	3.1.2000.7.0.3.0-79
	org.apache.hive.hcatalog	hive-hcatalog-streaming	3.1.2000.7.0.3.0-79
	org.apache.hive.shims	hive-shims-0.20	3.1.2000.7.0.3.0-79
	org.apache.hive.shims	hive-shims-0.23	3.1.2000.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.hive.shims	hive-shims-common	3.1.2000.7.0.3.0-79
	org.apache.hive.shims	hive-shims-scheduler	3.1.2000.7.0.3.0-79
Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector_2.11	1.0.0.7.0.3.0-79
Apache Kafka	org.apache.kafka	connect	2.3.0.7.0.3.0-79
	org.apache.kafka	connect-api	2.3.0.7.0.3.0-79
	org.apache.kafka	connect-basic-auth-extension	2.3.0.7.0.3.0-79
	org.apache.kafka	connect-file	2.3.0.7.0.3.0-79
	org.apache.kafka	connect-json	2.3.0.7.0.3.0-79
	org.apache.kafka	connect-runtime	2.3.0.7.0.3.0-79
	org.apache.kafka	connect-transforms	2.3.0.7.0.3.0-79
	org.apache.kafka	generator	2.3.0.7.0.3.0-79
	org.apache.kafka	jmh-benchmarks	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-clients	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-examples	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-log4j-appender	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-examples	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-scala_2.11	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-scala_2.12	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-test-utils	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-0100	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-0101	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-0102	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-0110	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-10	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-11	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-20	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-21	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-streams-upgrade-system-tests-22	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka-tools	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka_2.11	2.3.0.7.0.3.0-79
	org.apache.kafka	kafka_2.12	2.3.0.7.0.3.0-79
Knox	org.apache.knox	gateway-adapter	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-admin-ui	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-applications	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-cloud-bindings	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-demo-ldap	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-demo-ldap-launcher	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-discovery-ambari	1.3.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.knox	gateway-discovery-cm	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-docker	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-i18n	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-i18n-logging-log4j	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-i18n-logging-sl4j	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-ha	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-identity-assertion-common	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-identity-assertion-concat	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-identity-assertion-hadoop-groups	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-identity-assertion-pseudo	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-identity-assertion-regex	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-identity-assertion-switchcase	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-jersey	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-rewrite	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-rewrite-func-hostmap-static	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-rewrite-func-inbound-query-param	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-rewrite-func-service-registry	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-rewrite-step-encrypt-uri	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-rewrite-step-secure-query	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-authc-anon	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-authz-acls	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-clientcert	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-hadoopauth	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-jwt	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-pac4j	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-preauth	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-shiro	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-provider-security-webappsec	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-release	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-server	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-server-launcher	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-server-xforwarded-filter	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-admin	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-as	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-definitions	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-hashicorp-vault	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-hbase	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-health	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-hive	1.3.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.knox	gateway-service-idbroker	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-impala	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-jkg	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-knoxsso	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-knoxssout	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-knoxtoken	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-livy	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-nifi	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-nifi-registry	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-remoteconfig	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-rm	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-storm	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-test	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-tgs	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-vault	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-service-webhdfs	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-shell	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-shell-launcher	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-shell-release	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-shell-samples	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-spi	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-test	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-test-idbroker	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-test-release-utils	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-test-utils	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-util-common	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-util-configinjector	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-util-launcher	1.3.0.7.0.3.0-79
	org.apache.knox	gateway-util-urltemplate	1.3.0.7.0.3.0-79
	org.apache.knox	hadoop-examples	1.3.0.7.0.3.0-79
	org.apache.knox	knox-cli-launcher	1.3.0.7.0.3.0-79
	org.apache.knox	webhdfs-kerb-test	1.3.0.7.0.3.0-79
	org.apache.knox	webhdfs-test	1.3.0.7.0.3.0-79
Apache Kudu	org.apache.kudu	kudu-backup-tools	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-backup2_2.11	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-client	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-client-tools	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-flume-sink	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-hive	1.11.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.kudu	kudu-mapreduce	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-spark2-tools_2.11	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-spark2_2.11	1.11.0.7.0.3.0-79
	org.apache.kudu	kudu-test-utils	1.11.0.7.0.3.0-79
Livy	org.apache.livy	livy-api	0.5.0.7.0.3.0-79
	org.apache.livy	livy-client-common	0.5.0.7.0.3.0-79
	org.apache.livy	livy-client-http	0.5.0.7.0.3.0-79
	org.apache.livy	livy-core_2.11	0.5.0.7.0.3.0-79
	org.apache.livy	livy-examples	0.5.0.7.0.3.0-79
	org.apache.livy	livy-integration-test	0.5.0.7.0.3.0-79
	org.apache.livy	livy-repl_2.11	0.5.0.7.0.3.0-79
	org.apache.livy	livy-rsc	0.5.0.7.0.3.0-79
	org.apache.livy	livy-scala-api_2.11	0.5.0.7.0.3.0-79
	org.apache.livy	livy-server	0.5.0.7.0.3.0-79
	org.apache.livy	livy-test-lib	0.5.0.7.0.3.0-79
Lucene	org.apache.lucene	lucene-analyzers-common	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-icu	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-kuromoji	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-morfologik	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-nori	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-opennlp	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-phonetic	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-smartcn	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-stempel	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-analyzers-uima	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-backward-codecs	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-benchmark	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-classification	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-codecs	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-core	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-demo	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-expressions	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-facet	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-grouping	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-highlighter	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-join	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-memory	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-misc	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-queries	7.4.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.lucene	lucene-queryparser	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-replicator	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-sandbox	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-spatial	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-spatial-extras	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-spatial3d	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-suggest	7.4.0.7.0.3.0-79
	org.apache.lucene	lucene-test-framework	7.4.0.7.0.3.0-79
Apache Oozie	org.apache.oozie	oozie-client	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-core	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-distro	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-examples	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-fluent-job-api	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-fluent-job-client	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-server	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-distcp	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-git	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-hcatalog	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-hive	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-hive2	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-oozie	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-spark	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-sqoop	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-sharelib-streaming	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-tools	5.1.0.7.0.3.0-79
	org.apache.oozie	oozie-zookeeper-security-tests	5.1.0.7.0.3.0-79
	org.apache.oozie.test	oozie-mini	5.1.0.7.0.3.0-79
ORC	org.apache.orc	orc-core	1.5.1.7.0.3.0-79
	org.apache.orc	orc-examples	1.5.1.7.0.3.0-79
	org.apache.orc	orc-mapreduce	1.5.1.7.0.3.0-79
	org.apache.orc	orc-shims	1.5.1.7.0.3.0-79
	org.apache.orc	orc-tools	1.5.1.7.0.3.0-79
Phoenix	org.apache.phoenix	phoenix-client	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-core	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-hive	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-load-balancer	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-pherf	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-pig	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-queryserver	5.0.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.phoenix	phoenix-queryserver-client	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-server	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-spark	5.0.0.7.0.3.0-79
	org.apache.phoenix	phoenix-tracing-webapp	5.0.0.7.0.3.0-79
Ranger	org.apache.ranger	conditions-enrichers	2.0.0.7.0.3.0-79
	org.apache.ranger	credentialbuilder	2.0.0.7.0.3.0-79
	org.apache.ranger	embeddedwebserver	2.0.0.7.0.3.0-79
	org.apache.ranger	jisql	2.0.0.7.0.3.0-79
	org.apache.ranger	ldapconfigcheck	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-atlas-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-atlas-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-hbase-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-hbase-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-hdfs-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-hdfs-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-hive-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-hive-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-kafka-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-kafka-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-kms	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-kms-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-kms-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-knox-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-knox-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-kylin-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-kylin-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-nifi-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-nifi-registry-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-ozone-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-ozone-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-plugin-classloader	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-plugins-audit	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-plugins-common	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-plugins-cred	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-plugins-installer	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-sampleapp-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-solr-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-solr-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-sqoop-plugin	2.0.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.ranger	ranger-sqoop-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-storm-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-storm-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-tagsync	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-tools	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-util	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-yarn-plugin	2.0.0.7.0.3.0-79
	org.apache.ranger	ranger-yarn-plugin-shim	2.0.0.7.0.3.0-79
	org.apache.ranger	sampleapp	2.0.0.7.0.3.0-79
	org.apache.ranger	unixauthclient	2.0.0.7.0.3.0-79
	org.apache.ranger	unixauthservice	2.0.0.7.0.3.0-79
	org.apache.ranger	unixusersync	2.0.0.7.0.3.0-79
Ratis	org.apache.ratis	ratis-client	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-common	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-examples	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-grpc	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-hadoop	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-logservice	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-metrics	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-netty	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-proto	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-replicated-map	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-resource-bundle	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-server	0.4.0.7.0.3.0-79
	org.apache.ratis	ratis-test	0.4.0.7.0.3.0-79
Apache Solr	org.apache.solr	solr-analysis-extras	7.4.0.7.0.3.0-79
	org.apache.solr	solr-analytics	7.4.0.7.0.3.0-79
	org.apache.solr	solr-cell	7.4.0.7.0.3.0-79
	org.apache.solr	solr-clustering	7.4.0.7.0.3.0-79
	org.apache.solr	solr-core	7.4.0.7.0.3.0-79
	org.apache.solr	solr-dataimporthandler	7.4.0.7.0.3.0-79
	org.apache.solr	solr-dataimporthandler-extras	7.4.0.7.0.3.0-79
	org.apache.solr	solr-jetty-customizations	7.4.0.7.0.3.0-79
	org.apache.solr	solr-langid	7.4.0.7.0.3.0-79
	org.apache.solr	solr-ltr	7.4.0.7.0.3.0-79
	org.apache.solr	solr-prometheus-exporter	7.4.0.7.0.3.0-79
	org.apache.solr	solr-security-util	7.4.0.7.0.3.0-79
	org.apache.solr	solr-solrj	7.4.0.7.0.3.0-79
	org.apache.solr	solr-test-framework	7.4.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.solr	solr-uima	7.4.0.7.0.3.0-79
	org.apache.solr	solr-velocity	7.4.0.7.0.3.0-79
Apache Spark	org.apache.spark	spark-avro_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-catalyst_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-core_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-graphx_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-hadoop-cloud_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-hive-thriftserver_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-hive_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-kubernetes_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-kvstore_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-launcher_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-mllib-local_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-mllib_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-network-common_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-network-shuffle_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-network-yarn_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-repl_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-sketch_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-sql-kafka-0-10_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-sql_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming-flume-assembly_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming-flume-sink_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming-flume_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming-kafka-0-10_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming-kafka-0-8-assembly_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming-kafka-0-8_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-streaming_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-tags_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-unsafe_2.11	2.4.0.7.0.3.0-79
	org.apache.spark	spark-yarn_2.11	2.4.0.7.0.3.0-79
Apache Sqoop	org.apache.sqoop	sqoop	1.4.7.7.0.3.0-79
	org.apache.sqoop	sqoop-test	1.4.7.7.0.3.0-79
Tez	org.apache.tez	hadoop-shim	0.9.1.7.0.3.0-79
	org.apache.tez	hadoop-shim-2.8	0.9.1.7.0.3.0-79
	org.apache.tez	tez-api	0.9.1.7.0.3.0-79
	org.apache.tez	tez-aux-services	0.9.1.7.0.3.0-79
	org.apache.tez	tez-common	0.9.1.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.tez	tez-dag	0.9.1.7.0.3.0-79
	org.apache.tez	tez-examples	0.9.1.7.0.3.0-79
	org.apache.tez	tez-ext-service-tests	0.9.1.7.0.3.0-79
	org.apache.tez	tez-history-parser	0.9.1.7.0.3.0-79
	org.apache.tez	tez-javadoc-tools	0.9.1.7.0.3.0-79
	org.apache.tez	tez-job-analyzer	0.9.1.7.0.3.0-79
	org.apache.tez	tez-mapreduce	0.9.1.7.0.3.0-79
	org.apache.tez	tez-protobuf-history-plugin	0.9.1.7.0.3.0-79
	org.apache.tez	tez-runtime-internals	0.9.1.7.0.3.0-79
	org.apache.tez	tez-runtime-library	0.9.1.7.0.3.0-79
	org.apache.tez	tez-tests	0.9.1.7.0.3.0-79
	org.apache.tez	tez-yarn-timeline-cache-plugin	0.9.1.7.0.3.0-79
	org.apache.tez	tez-yarn-timeline-history	0.9.1.7.0.3.0-79
	org.apache.tez	tez-yarn-timeline-history-with-acls	0.9.1.7.0.3.0-79
	org.apache.tez	tez-yarn-timeline-history-with-fs	0.9.1.7.0.3.0-79
Zeppelin	org.apache.zeppelin	sap	0.8.0.7.0.3.0-79
	org.apache.zeppelin	spark-interpreter	0.8.0.7.0.3.0-79
	org.apache.zeppelin	spark-scala-2.11	0.8.0.7.0.3.0-79
	org.apache.zeppelin	spark-shims	0.8.0.7.0.3.0-79
	org.apache.zeppelin	spark2-shims	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-alluxio	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-angular	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-bigquery	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-cassandra_2.10	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-display	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-elasticsearch	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-file	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-flink_2.10	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-groovy	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-hbase	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-ignite_2.10	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-interpreter	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-jdbc	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-jupyter	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-kylin	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-lens	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-livy	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-markdown	0.8.0.7.0.3.0-79
	org.apache.zeppelin	zeppelin-neo4j	0.8.0.7.0.3.0-79

Project	groupId	artifactId	version
	org.apache.zookeeper	zookeeper	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-client-c	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-contrib-loggraph	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-contrib-rest	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-docs	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-jute	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-recipes-election	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-recipes-lock	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-recipes-queue	3.5.5.7.0.3.0-79
Apache ZooKeeper	org.apache.zookeeper	zookeeper	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-client-c	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-contrib-loggraph	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-contrib-rest	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-docs	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-jute	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-recipes-election	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-recipes-lock	3.5.5.7.0.3.0-79
	org.apache.zookeeper	zookeeper-recipes-queue	3.5.5.7.0.3.0-79

What's New In Cloudera Runtime 7.0.3

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

Related Information

[Overview](#)

What's New in Apache Atlas

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

Associated Classification Counts and Search Links in Statistics

Atlas Statistics page now includes a list of the classifications that have been associated with entities. For each classification, Atlas shows a count of the associated entities; clicking the count opens search results for a search on that classification.

See [Apache Atlas Statistics reference](#)

New Atlas Metrics for Cloudera Manager Charts

Atlas now provides a number of metrics that you can use to create time-series charts in Cloudera Manager. Add these charts to a dashboard in Cloudera Manager to see Atlas resource use at a glance.

Metrics include the rate of entity creation, update, and deletion and the size of Java heap used by Atlas.

See [Charting Time-Series Data](#)

What's new in DAS

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

What's New in Apache HBase

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

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

What's New in Apache Hadoop HDFS

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

There are no new Apache Hive features in this release.

What's New in Hue

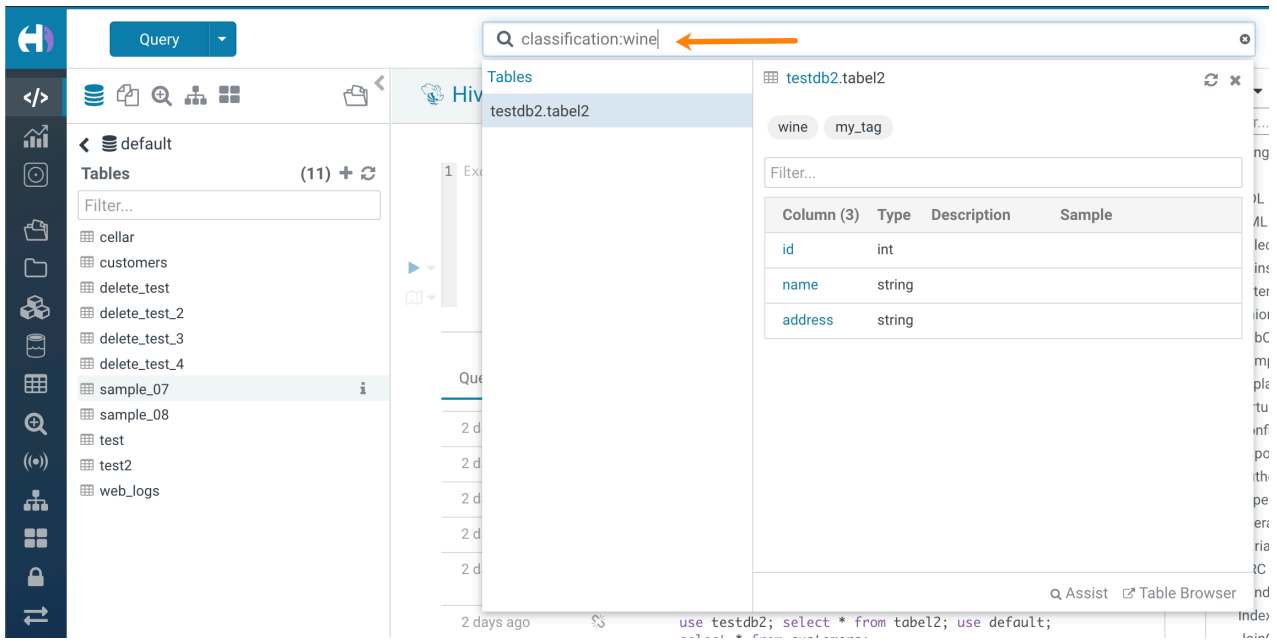
This topic lists new features for Hue in this release of Cloudera Runtime.

Support for Hive on Tez

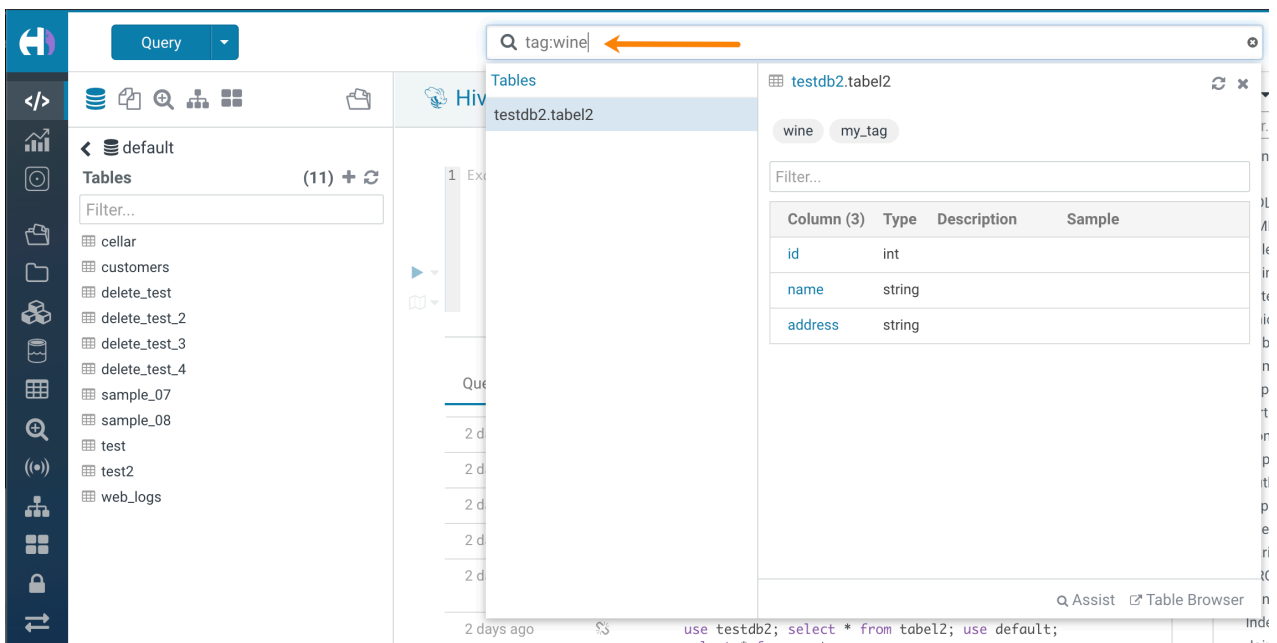
Hue in Cloudera Runtime now supports Hive using Tez as its execution engine.

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:

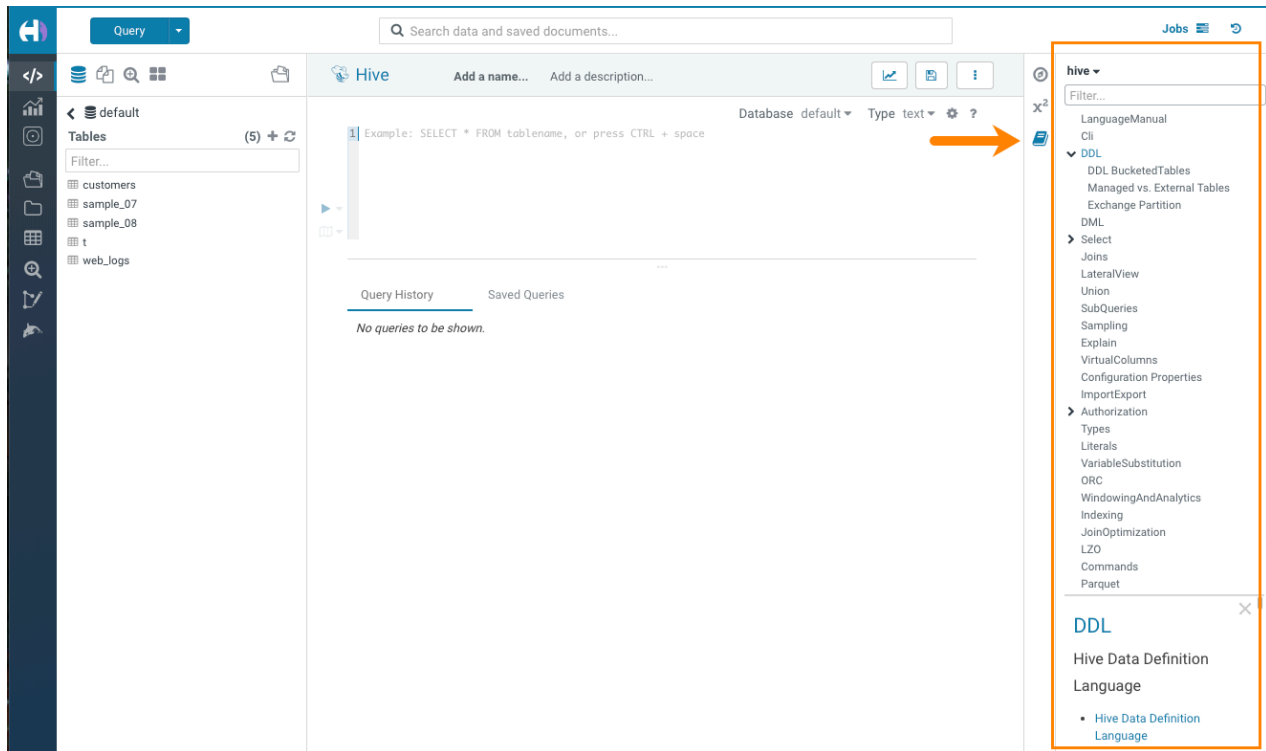


You can also locate tables by searching for Atlas classifications by specifying the tag search term:



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.

Increased Compatibility with Components in Cloudera Data Platform

Impala is integrated with the following components:

- Apache Ranger: Use Apache Ranger to manage authorization in Impala. See [Impala Authorization](#) for details.
- Apache Atlas: Use Apache Atlas to manage data governance in Impala. See [Atlas Metadata Collection Overview](#)
- Hive 3

Parquet Page Index

To improve performance when using Parquet files, Impala can now write page indexes in Parquet files and use those indexes to skip pages for the faster scan.

See [Using Parquet in Impala](#) for more information.

New Compression Supported for Parquet Files

Impala supports the Lz4 compression for the tables in the Parquet file format.

Lz4 is a lossless compression algorithm providing extremely fast and scalable compression and decompression.

Default File Format Changed to Parquet

When you create a table, the default format for that table data is now Parquet unless the STORED AS clause is specified.

For backward compatibility, you can use the `DEFAULT_FILE_FORMAT` query option to set the default file format to the previous default, such as text or other formats.

Metadata Performance Improvements

The following features for improving metadata performance are enabled by default in this release:

- Incremental stats are now compressed in memory in catalogd, reducing memory footprint in catalogd.
- impalad coordinators fetch incremental stats from catalogd on-demand, reducing the memory footprint and the network requirements for broadcasting metadata.
- Time-based and memory-based automatic invalidation of metadata to keep the size of metadata bounded and to reduce the chances of catalogdcache running out of memory.
- Automatic invalidation of metadata

With automatic metadata management enabled, you no longer have to issue `INVALIDATE / REFRESH` in a number of conditions.

See [Impala Metadata Management](#) for the information on the above features.

Support for Kudu Integrated with Hive Metastore

Kudu is now integrated with Hive Metastore (HMS), and from Impala, you can create, update, delete, and query the tables in the Kudu services that are integrated with HMS.

See [Using Kudu with Impala](#) for information on using Kudu tables in Impala.

DATE Data Type and Functions

You can use the new `DATE` data type to describe particular year/month/day values.

This initial `DATE` type supports the Text, Parquet, Avro, and HBASE file formats.

Most of the built-in functions for `TIMESTAMP` now allow the `DATE` type arguments, as well.

The support of `DATE` data type includes the following features:

- `DATE` type column as a partitioning key column
- `DATE` literal
- Implicit casting between `DATE` and other types, namely, `STRING` and `TIMESTAMP`

See [DATE Data Type](#) and [Impala Date and Time Functions](#) for using the `DATE` type.

Support for Hive Insert-Only Transactional Tables

Impala added the support to create, drop, truncate, query, and insert into insert-only transactional tables.

By default, Impala creates an insert-only transactional table when you issue the `CREATE TABLE` statement.

Use the Hive compaction to compact small files to improve the performance and scalability of metadata in transactional tables.

See [Impala Transactions](#) for more information.

Server-side Spooling of Query Results

You can use the `SPOOL_QUERY_RESULTS` query option to control how query results are returned to the client.

By default, when a client fetches a set of query results, the next set of results are fetched in batches until all the result rows are produced. If a client issues a query without fetching all the results, the query fragments continue to hold on to the resources until the query is canceled and unregistered, potentially tying up resources and causing other queries to wait in admission control.

When the query result spooling feature is enabled, the result sets of queries are eagerly fetched and buffered until they are read by the client, and resources are freed up for other queries.

See [Spooling Impala Query Results](#) for the new feature and the query options.

HiveServer2 HTTP Connection for Clients

Now client applications can connect to Impala over HTTP via HiveServer2 with the option to use the Kerberos SPNEGO and LDAP for authentication. See [Impala Clients](#) for details.

Cookie-based Authentication

Starting in this version, Impala supports cookies for authentication when clients connect via HiveServer2 over HTTP.

You can use the `--max_cookie_lifetime_s` startup flag to:

- Disable the use of cookies
- Control how long generated cookies are valid for

See [Impala Clients](#) for more information.

Support of Kerberos Authentication in Impala Web UI

Starting in this release, you can configure Kerberos authentication in Cloudera Manager to secure the debug Web UI pages for Impala Daemon, Catalog Server, and StateStore.

See [Configuring Impala Web UI](#) for the steps to enable Kerberos authentication for the Impala Web UI pages.

Object Ownership Support

Object ownership for tables, views, and databases is enabled by default in Impala. When you create a database, a table, or a view, as the owner of that object, you implicitly have the privileges on the object. The privileges that owners have are specified in Ranger on the special user, {OWNER}.

The {OWNER} user must be defined in Ranger for the object ownership privileges work in Impala.

See [Authorization](#) for details.

Graceful Shutdown of Impala Daemons

You can perform a graceful shutdown of Impala Daemons in Cloudera Manager.

When you initiate a shutdown process for an Impala Daemon, the Impala Daemon notifies other Impala Daemons that it is shutting down, waits for a grace period, then shuts itself down once no more queries or fragments are executing on that daemon or when the configurable deadline is reached.

See [Graceful Shutdown](#) for the steps.

New Built-in Functions for Fuzzy Matching of Strings

Use the new Jaro or Jaro-Winkler functions to perform fuzzy matches on relatively short strings, e.g. to scrub user inputs of names against the records in the database.

- JARO_DISTANCE, JARO_DST
- JARO_SIMILARITY, JARO_SIM
- JARO_WINKLER_DISTANCE, JW_DST
- JARO_WINKLER_SIMILARITY, JW_SIM

See [Impala String Functions](#) for details.

New Built-in Function to Process JSON Objects

The GET_JSON_OBJECT() function extracts JSON object from a string, based on the path specified as an input argument, and returns the extracted JSON object.

See [Impala Miscellaneous Functions](#).

Capacity Quota for Scratch Disks

When configuring scratch space for intermediate files used in large sorts, joins, aggregations, or analytic function operations, use the `##scratch_dirs` startup flag to optionally specify a capacity quota per scratch directory, e.g., `##scratch_dirs=/dir1:5MB,/dir2`.

See [Impala and HDFS](#) for details.

S3 Supported for Remote File Handle Cache

Impala can now cache remote HDFS file handles for the tables that store their data in Amazon S3 cloud storage.

See [Impala Scalability Considerations](#) for the information on remote file handle cache.

Data Cache for Remote Reads

You can execute queries faster on multi-cluster HDFS environments and on object store environments as Impala now caches data for non-local reads (e.g. S3, ABFS, ADLS) on local storage.

The data cache is enabled with the `--data_cache` startup flag.

See [Impala Remote Data Cache](#) for the information and steps to enable remote data cache.

Query Option for Disabling HBase Row Estimation

During query plan generation, Impala samples underlying HBase tables to estimate row count and row size, but the sampling process can negatively impact the planning time. To alleviate the issue, when the HBase table stats do not change much in a short time, disable the sampling with the `DISABLE_HBASE_NUM_ROWS_ESTIMATE` query option so that the Impala planner falls back to using Hive Metastore (HMS) table stats instead.

See [Impala Query Options](#).

Query Option for Controlling Size of Parquet Splits on Non-block Stores

To optimize query performance, Impala planner uses the value of the `fs.s3a.block.size` startup flag when calculating the split size on non-block based stores, e.g. S3, ADLS, etc. Starting in this release, Impala planner uses the `PARQUET_OBJECT_STORE_SPLIT_SIZE` query option to get the Parquet file format specific split size.

For Parquet files, the `fs.s3a.block.size` startup flag is no longer used.

The default value of the `PARQUET_OBJECT_STORE_SPLIT_SIZE` query option is 256 MB.

See [Impala with Amazon S3](#) for tuning Impala query performance for S3.

Scalable Pool Configuration in Admission Controller

To offer more dynamic and flexible resource management, Impala supports the new configuration parameters that scale with the number of executors. You can use the parameters to control the number of running queries, queued queries, and maximum amount of memory allocated for Impala resource pools.

See [Impala Admission Control](#) for the information about the new parameters and using them for admission control.

Query Profile

The following metrics were added to the Query Profile output for better monitoring and troubleshooting of query performance.

- Network I/O throughput
- System disk I/O throughput

See [Impala Query Profile](#) for generating and reading query profile.

Query Profile Exported to JSON

On the Query Details page of Impala Daemon Web UI, you have a new option, in addition to the existing Thrift and Text formats, to export the query profile output in the JSON format.

See [Impala Daemon Web UI](#) for generating JSON query profile outputs in Web UI.

What's New in Apache Kafka

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

Rebase on Apache Kafka 2.3.0

Kafka shipped with this version of Cloudera Runtime is based on Apache Kafka 2.3.0. For more information, see [Apache Kafka Notable Changes](#) and [Apache Kafka Release Notes](#) in the upstream documentation.

LDAP authentication support

You can now configure Kafka to authenticate clients using LDAP. For more information, see [LDAP Authentication](#)

Ranger authorization support

Ranger support for Kafka is added. You can now use Ranger to provide authorization for Kafka. For more information, see [Using Ranger to Provide Authorization in CDP](#) as well as the documentation on [Kafka Authorization with Ranger](#).

Enable Delegation Tokens is set to true by default

The Enable Delegation Tokens property that enables and disables delegation token based authentication is set to true by default. Note however, that this does not by default enable the feature itself as delegation tokens require Kerberos authentication to function.

For more information, see [Enable or Disable Authentication with Delegation Tokens](#).

New metric for monitoring garbage collector runs

A new metric called `kafka_jvm_gc_runs` is added to the Kafka service. This metric enables users to monitor the number of garbage collector runs performed on each broker.

What's New in Apache Kudu

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

Support for EC-based ciphers

Kudu supports EC-based ciphers in the web UI. Previously, both kudu-tserver and kudu-master failed to start if supplied with server certificates based on ciphers that squeasel does not support.

Support for putting tablet servers in to maintenance mode

Kudu supports putting tablet servers into maintenance. In this mode, the tablet server's replicas are not re-replicated if they fail. Re-replication for the remaining, under-replicated tablets is triggered only when you exit from the maintenance mode. The `kudu tserver state enter_maintenance` and the `kudu tserver state exit_maintenance` tools have been added to orchestrate the tablet server maintenance. The `kudu tserver list` tool has been amended with a "state" column option to display the current state of each tablet server.

Built-in NTP client maintains internal time

Kudu has a built-in NTP client which maintains an internal time that is used to generate the HybridTime timestamps. If you enable the NTP client, the system clock synchronization for the nodes running Kudu is no longer necessary. This is useful for containerized deployments and in other cases when it is troublesome to maintain a properly configured system NTP service at each node of a Kudu cluster. The list of NTP servers to synchronize against is specified with the `--builtin_ntp_servers` flag. By default, the Kudu masters and the tablet servers use public servers hosted by the NTP Pool project. To use the built-in NTP client, set the `--time_source=builtin` flag and reconfigure the `--builtin_ntp_servers` flag if necessary.

Support for aggregated table statistics for the Kudu clients

Aggregated table statistics are now available to the Kudu clients through the `KuduClient.getTableStatistics()` and the `KuduTable.getTableStatistics()` methods in the Kudu Java client, and by using the `KuduClient.GetTableStatistics()` method in the Kudu C++ client. This allows for various query optimizations.

For example, Spark now uses the aggregated table statistics to perform join optimizations. The statistics are available via the API of both the C++ and the Java Kudu clients. In addition, per-table statistics are available via the kudu table statistics CLI tool. The statistics are also available via the master's Web UI at the `master:8051/metrics` and the `master:8051/table?id=<uuid>` URIs.

New operations using the Kudu CLI

The Kudu CLI supports the following new operations:

- Altering table columns: The following, newly introduced sub-commands allow you to alter a column of the specified table:
 - `kudu table column_set_default`
 - `kudu table column_remove_default`
 - `kudu table column_set_compression`
 - `kudu table column_set_encoding`
 - `kudu table column_set_block_size`
- Dropping table columns: The `kudu table delete_column` sub-command allows you to drop a column of the specified table.
- Getting and setting extra configuration properties: The `kudu table get_extra_configs` and the `kudu table set_extra_config` sub-commands allow you to get and set the extra extra configuration properties for a table respectively.
- Creating and dropping range partitions: The `kudu table add_range_partition` and the `kudu table drop_range_partition` sub-commands allow you to create and drop the range partitions for a table respectively.

Optimizations and improvements

- Tablet servers now expand a tablet's data directory group with available healthy directories when all directories of the group are full.
- For scan operations that are run with the `CLOSEST_REPLICA` selection mode, the Kudu Java client now picks a random available replica in case no replica is located at the same node with the client that initiated the scan operation. This helps to spread the load generated by multiple scan requests to the same tablet among all available replicas. In older versions of Kudu, all such scan requests would end up fetching data from the same tablet replica.
- The tablet servers now consider the available disk space when choosing a set of data directories for a tablet's data directory group, and when deciding in which data directory a new block should be written.
- The tablet servers reject any individual write operations that violate the schema constraints in a batch of write operations that are received from a client. The previous behavior was to reject the whole batch of write operations if a violation of the schema constraints is detected even for a single row.
- Kudu RPC now enables TCP keepalive for all outbound connections for faster detection of the no-longer-reachable nodes.
- The memory reserved by `tcMalloc` is now released to the OS periodically to avoid any potential OOM issues in case of read-only workloads.

- The evaluation of predicates on columns of primitive types and NULL or NOT NULL predicates has been optimized to leverage SIMD instructions.

What's New in Apache Oozie

This topic lists 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 release of Cloudera Runtime introduces Apache Hadoop Ozone, a scalable, redundant, and distributed object store for big data applications.



Important:

Apache Hadoop Ozone in CDP is a technical preview and considered under development. Do not use this in your production systems. If you have feedback, contact Support by logging a case on our Cloudera Support Portal at <https://my.cloudera.com/support.html>. Technical preview features are not guaranteed troubleshooting and fixes.

For more information about Ozone, see [Introduction to Ozone](#).

What's New in Apache Phoenix

This release of Cloudera Runtime introduces Apache Phoenix, a massively parallel, relational database engine supporting OLTP use cases using Apache HBase as its backing store.

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

What's New in Apache Ranger

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

Reduced latency for Ranger policy enforcement

Apache Ranger now performs incremental sync of Atlas tags and incremental policy downloads into Ranger plugins, which reduces latency of Ranger policy enforcement as compared to the previous "flush and fill" model.

API support for Sentry policy migration to Ranger

New Ranger bulk import/creation APIs for Ranger policies facilitate migration of Sentry authorization policies into Ranger.

Enhancements to support replication of wildcard and tag-based policies: DLM/BDR use cases

Semantics for wildcard and tag-based policies are preserved when authorization policies and associated data are copied between source and target clusters. This enhancement supports DLM (Data Lifecycle Management) and BDR (Backup and Data Recovery) use cases.

Solr plugin enhancements to support document-level authorization (Sentry Feature Parity)

To support Sentry feature parity, the Ranger Solr plugin now supports document-level authorization.

Enhancements in Ranger plugins to support Ranger HA without load-balancer

You can now deploy Ranger in HA mode in CDP without requiring additional external load balancer hardware or software.

Java 11 support for Ranger

Apache Ranger now supports Java 11.

Ranger support for CM monitoring and metrics

Ranger monitoring and metrics are now accessible via Cloudera Manager.

Ranger authorization for Ozone (Tech Preview)

Ranger authorization policies can now be used with Ozone storage. Ozone support is a technical preview feature.

What's New in Cloudera Search

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

As of this release, Cloudera Search is available with Cloudera Runtime.

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

What's New in Apache Spark

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

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

What's New in Apache Hadoop YARN

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

Queue Manager

YARN Queue Manager is the queue management graphical user interface for Apache Hadoop YARN Capacity Scheduler. You can use YARN Queue Manager UI to manage your cluster capacity using queues to balance resource requirements of multiple applications from various users. Using YARN Queue Manager UI, you can set scheduler level properties and queue level properties. You can also view, sort, search, and filter queues using the YARN Queue Manager UI.

For more information about Queue Manager, see [Allocating Resources Using YARN Queue Manager UI](#)

More reading

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

What's New in Apache ZooKeeper

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

Enable 4LW telnet commands in ZooKeeper configuration using CM

You can configure the `4lw.commands.whitelist` property using Cloudera Manager. By default the `wchc` and `wchp` commands are not enabled because of their known DOS vulnerability. For more information, see [Configure four-letter-words commands in ZooKeeper using Cloudera Manager](#).

New diagnostics REST interface called AdminServer

AdminServer is a new diagnostics REST interface that provides simple diagnostics and metrics through HTTP. The embedded admin server is optional and disabled by default. You can enable it and change the server port using Cloudera Manager. For more information, see [Enable the AdminServer](#).

Known Issues In Cloudera Runtime 7.0.3

Summary of known issues in this release.

Related Information

[Overview](#)

Known Issues in Apache Atlas

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

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

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

Workaround: None.

Cloudera JIRA: CDPD-7982

Lineage info for HBase not available in a TLS-enabled environment

Atlas will not collect metadata from HBase (namespaces, tables, or column families) when Cloudera Manager is configured for TLS/SSL. This issue has no effect on HBase functionality, however, tag-based authorization through Ranger policies will not be available for HBase.

Workaround: None.

Cloudera JIRA: OPSAPS-53788

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

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.

Workaround: None.

Cloudera JIRA: OPSAPS-51224

Unexpected Results from Atlas Advanced Search Query with Underscore Character

An Atlas Advanced Search with an equal operator (=) and a like operator where the like object includes an underscore does not return correct results. For example, the following search does not return expected results:

db.name : "=" finance and QualifiedName : "like" *finance_09*

However, each part of the query run separately or the same query without the underscore return results as expected.

Workaround: None.

Cloudera JIRA: CDPD-5314

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

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 Apache Accumulo

This topic describes known issues and workarounds for using Accumulo in this release of Cloudera Runtime. **Apache Accumulo is currently not supported in this version of Cloudera Runtime. Although you can access Accumulo from the command-line interface, you must not use this component in production because Cloudera does not support it.**

Known Issues in DAS

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

- The task logs for a particular task may not be available in the task swimlane. And the zip file generated by download logs artifact may not have task logs, but instead contain an error-reports.json file with the error log of the download failures.
- You may not see any data for a report for any new queries that you run. This can happen especially for the last one day's report.

Workaround:

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

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

3. Start the DAS Event Processor.
- 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).
- You are unable to see databases or the query editor on the **Compose** page even after following the troubleshooting procedure. This can be caused due to incorrect hive service name in ZooKeeper.
- 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.
- 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.

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 HBase

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

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

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

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

HDFS encryption with HBase

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

Workaround: N/A

AccessController postOperation problems in asynchronous operations

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

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

Workaround: N/A

Apache Issue: [HBASE-6992](#)

Bulk loading into HBase configured to use cloud storage (S3) is currently not supported

When HBase is configured to use cloud storage (S3), bulk loading data into HBase fails with the following error: 401 Authentication required.

Workaround: N/A

Apache Issue: N/A

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

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

Workaround: N/A

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 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.

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

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

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](#)

Known Issues in Apache Hive

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

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.

CDPD-6030: External table with custom TDE zone cannot be accessed.

Workaround: Use default location for external tables and use TDE for default location.

Apache JIRA: none

CDPD-5962: A schema evolution that replaces a date type with a string type fails.

Workaround: none

Apache JIRA: none

Technical Service Bulletins

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 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.

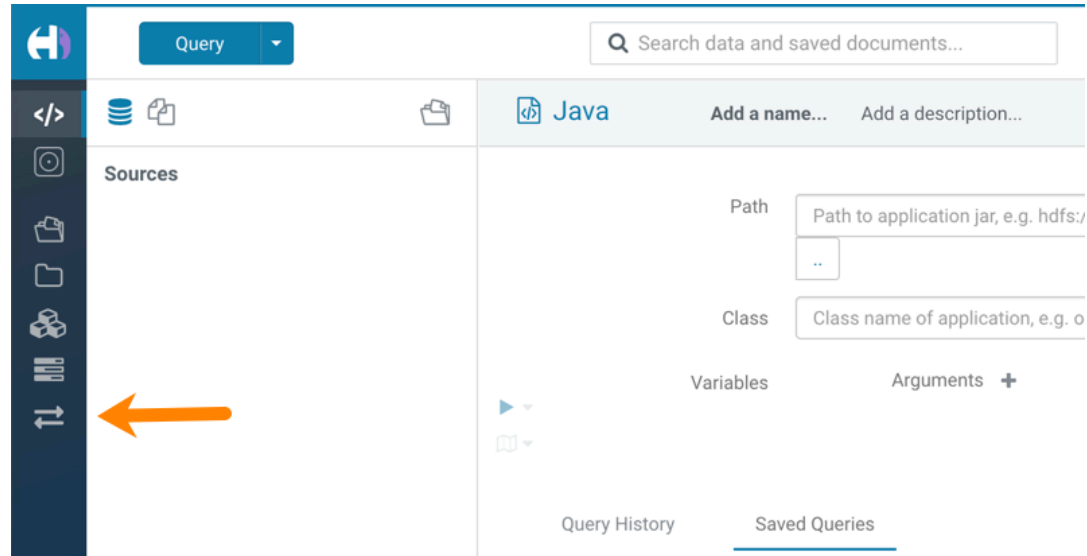
Hive and Impala query editors do not work with TLS 1.2

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:



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.

For the known issues and workarounds in Impala, see the [Impala Known Issues in CDP](#) Knowledge Base article.

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

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

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

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

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

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

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

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

whether a major compaction was run on the table and on whether the old files compacted during a minor compaction have been deleted.

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

Minor compaction is not initiated automatically by Hive Metastore (HMS) or any other CDP (Cloudera Data Platform) component, meaning that this issue can only occur if minor 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 and unsupported features for using Kafka in this release of Cloudera Runtime.

Known Issues

Topics created with the kafka-topics tool might not be secured when the deprecated --zookeeper option is used

Topics that are created and deleted via Kafka are secured (for example, auto created topics). 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. Anyone with access to ZooKeeper can create and delete topics. They will not be able to describe, read, or write to the topics even if they can create them.

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

Significant performance degradation can occur when SSL is enabled. The impact varies depending on your CPU, JVM version, 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.

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.
- Kafka Connect is not supported. NiFi is a proven solution for batch and real time data loading that complement Kafka's message broker capability. For more information, see [Cloudera Flow Management](#).
- The Kafka default authorizer is not supported. This includes setting ACLs and all related APIs, broker functionality, and command-line tools.

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 Ranger
 - 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 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.

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

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

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

Unsupported Feature

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

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

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

Known Issues in Cloudera Search

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

Known Issues

Splitshard of HDFS index checks local filesystem and fails

When performing a shard split on an index that is stored on HDFS, SplitShardCmd still evaluates free disk space on the local file system of the server where Solr is installed. This may cause the command to fail, perceiving that there is no adequate disk space to perform the shard split.

Workaround: None

Apache Issue: SOLR-14251

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

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

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

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

if there are no collections present in Solr.

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

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.

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

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.

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

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

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

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

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

CrunchIndexerTool which includes Spark indexer requires specific input file format specifications

If the `--input-file-format` option is specified with `CrunchIndexerTool`, then its argument must be text, avro, or avroParquet, rather than a fully qualified class name.

Workaround: None.

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

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

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`.

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

The `MapReduceIndexerTool` and the `HBaseMapReduceIndexerTool` can be used with a Managed Schema created via NRT indexing of documents or via the Solr Schema API. However, neither tool supports adding fields automatically to the schema during ingest.

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.

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

The Browse and Spell Request Handlers require certain fields to be present in the schema. Since those fields cannot be guaranteed to exist in a Schemaless setup, the Browse and Spell Request Handlers are not enabled by default.

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.

Enabling blockcache writing may result in unusable indexes

It is possible to create indexes with `solr.hdfs.blockcache.write.enabled` set to `true`. Such indexes may appear corrupt to readers, and reading these indexes may irrecoverably corrupt indexes. Blockcache writing is disabled by default.

Workaround: None.

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

Users who are not authorized to use the Solr Admin UI are not given a page explaining that access is denied to them, instead receive a web page that never finishes loading.

Workaround: None.

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

Repeatedly running the MapReduceIndexerTool on the same set of input files can result in duplicate entries in the Solr collection. This occurs because the tool can only insert documents and cannot update or delete existing Solr documents. This issue does not apply to the HBaseMapReduceIndexerTool unless it is run with more than zero reducers.

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

Deleting collections might fail if hosts are unavailable

It is possible to delete a collection when hosts that host some of the collection are unavailable. After such a deletion, if the previously unavailable hosts are brought back online, the deleted collection may be restored.

Workaround: Ensure all hosts are online before deleting collections.

Saving search results is not supported

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

Workaround: None.

HDFS Federation is not supported

Cloudera Search does not support HDFS Federation.

Workaround: None.

Collection state goes down after Solr SSL

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

Workaround: Recreate the collection after enabling TLS.

Unsupported Features

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

- [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).
- [Logging Slow Queries](#)

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 Spark

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

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

Workaround: Use a different compression algorithm.

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

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

Workaround: Create the view using Hive or Impala.

Known Issues for Apache Sqoop

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

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

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.

CDPD-2936: Application logs are not accessible in WebUI2 or Cloudera Manager

Running Containers Logs from NodeManager local directory cannot be accessed either in Cloudera Manager or in WebUI2 due to log aggregation.

Workaround: Use the YARN log CLI to access application logs. For example:

```
yarn logs -applicationId <Application ID>
```

Apache Issue: [YARN-9725](#)

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

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

Workaround: You can manually add the required environment variables to the whitelist using Cloudera Manager.

1. In Cloudera Manager, select the YARN service.

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

COMPX-1445: Queue Manager operations are failing when Queue Manager is installed separately from YARN

If Queue Manager is not selected during YARN installation, Queue Manager operation are failing. Queue Manager says 0 queues are configured and several failures are present. That is because ZooKeeper configuration store is not enabled.

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 YARM 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-8687: Missing access check for getAppAttempts

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

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

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

Limitations

Capacity Scheduler

- As Capacity Scheduler is the default scheduler, the Dynamic Resource Pool User Interface is not available by default.
- Capacity Scheduler can be configured only through safety-valves in Cloudera Manager.

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
- New Aggregated Log File Format
- Node Labels
- Rolling Log Aggregation
- YARN WebUI v2
- Docker on YARN (DockerContainerExecutor)
- Moving jobs between queues
- Dynamic Resource Pools
- Upgrading from Cloudera Data Hub (CDH) or Hortonworks Data Platform (HDP) to Cloudera Data Platform - Data Center (CDP-DC) 7.0

CDP-DC 7.0 release only supports Capacity Scheduler for your Hadoop YARN clusters. However, you cannot upgrade to the CDP-DC 7.0 release from your existing CDH or HDP clusters. Therefore, in this release, we do not support any sort of migrations from CDH or HDP clusters and from Fair Scheduler to Capacity Scheduler. If you wish to migrate from Fair Scheduler to Capacity Scheduler, you must first upgrade to Cloudera Data Platform – Data Center 7.1 release and migrate from Fair Scheduler to Capacity Scheduler. For more information, see [Migrating from Fair Scheduler to Capacity Scheduler](#).

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](#)

Cloudera Runtime Component Versions

List of the component versions for Cloudera Runtime. You must use this as a reference for Cloudera Runtime and Encryption 7.0.3 components.

Table 1: Components included in the Cloudera Runtime parcel

Component	Version
Apache Atlas	2.0.0
Apache Avro	1.8.2
Apache Hadoop	3.1.1
Apache HBase	2.2.2
HBase Indexer	1.5.0
Apache Hive	3.1.2
Hue	4.5.0
Apache Impala	3.2.0
Apache Kafka	2.3.0
Apache Kudu	1.11.0
Apache Oozie	5.1.0
Apache ORC	1.5.1
Apache Ozone	0.4.0 (Technical Preview)
Apache Parquet	1.10.99
Apache Phoenix	5.0.0
Apache Ranger	2.0.0
Apache Solr	7.4.0
Apache Spark	2.4.4
Apache Sqoop	1.4.7
Apache Tez	0.9.1
Apache Zookeeper	3.5.5
DAS	1.4.2

Table 2: Encryption components

Java KeyStore KMS	3.1
Key Trustee KMS	7.0
Key Trustee Server	7.0

Related Information

[Overview](#)

Updated HDP Components

Component Changes in CDP Data Center 7.x as compared to HDP

Component Changes in CDP Data Center 7.0

Flume, Storm, and Ambari have been removed.

- Flume workloads can be migrated to Cloudera Flow Management (CFM). CFM is a no-code data ingestion and management solution powered by Apache NiFi. Contact your Cloudera account team for more information about moving from Flume to CFM.
- Storm can be replaced with Cloudera Streaming Analytics (CSA) powered by Apache Flink. Contact your Cloudera account team for more information about moving from Storm to CSA.
- Cloudera Manager has replaced Ambari, but Ambari will be used as part of the upgrade process from HDP to CDP Data Center 7.1.

Related Information

[Overview](#)

Updated CDH Components

Component Changes in CDP Data Center 7.x as compared to CDH

Component Changes in CDP Data Center 7.0

Pig, Flume, Sentry, and Navigator have been removed.

- Pig can be replaced with Hive or Spark.
- Flume has been replaced with Cloudera Flow Management (CFM). CFM is a no-code data ingestion and management solution powered by Apache NiFi. Contact your Cloudera account team for more information about moving from Flume to CFM.
- Sentry has been replaced with Ranger. A Sentry-to-Ranger policy migration tool is available for CDP Data Center 7.1 and migrations will be supported when Replication Manager is used to migrate Hive tables from CDH to CDP.
- Navigator has been replaced with Atlas. Navigator lineage data is transferred to Atlas as part of the CDH to CDP Data Center upgrade process. Navigator audit data is not transferred to Atlas.