

Cloudera Runtime 7.2.17

Release Notes

Date published: 2023-06-20

Date modified:

CLOUdera

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

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

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

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

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

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

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

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

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

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

Contents

Overview.....	6
Cloudera Runtime Component Versions.....	6
Using the Cloudera Runtime Maven repository 7.2.17.....	7
Runtime 7.2.17.0-334.....	8
What's New In Cloudera Runtime 7.2.17.....	25
What's New in Apache Atlas.....	25
What's New in Cruise Control.....	26
Cruise Control Rebase Summary.....	26
What's new in Data Analytics Studio.....	27
What's New in Apache HBase.....	27
What's New in Apache Hive.....	28
What's New in Hue.....	29
What's new in Apache Iceberg.....	29
What's New in Apache Impala.....	30
What's New in Apache Kafka.....	32
What's New in Apache Knox.....	34
What's New in Apache Kudu.....	34
What's New in Apache Phoenix.....	35
What's New in Apache Ranger.....	35
What's New in Schema Registry.....	35
What's New in Apache Spark.....	36
What's New in Sqoop.....	36
What's new in Streams Messaging Manager.....	37
What's New in Streams Replication Manager.....	38
What's New in Apache Hadoop YARN and YARN Queue Manager.....	38
What's New in Apache ZooKeeper.....	39
Unaffected Components in this release.....	39
Fixed Issues In Cloudera Runtime 7.2.17.....	39
Fixed Issues in Atlas.....	39
Fixed Issues in Avro.....	43
Fixed Issues in Cloud Connectors.....	43
Fixed issues in Cruise Control.....	44
Fixed issues in Data Analytics Studio.....	44
Fixed Issues in Apache Hadoop.....	44
Fixed Issues in HBase.....	44
Fixed Issues in HDFS.....	47
Fixed Issues in Apache Hive.....	48
Fixed Issues in Hive Warehouse Connector.....	49
Fixed Issues in Hue.....	50
Fixed Issues in Apache Impala.....	50
Fixed Issues in Apache Kafka.....	52

Fixed Issues in Apache Knox.....	52
Fixed Issues in Apache Kudu.....	53
Fixed Issues in Apache Livy.....	54
Fixed Issues in Apache Oozie.....	54
Fixed Issues in Ozone.....	55
Fixed Issues in Phoenix.....	55
Fixed Issues in Parquet.....	57
Fixed Issues in Apache Ranger.....	57
Fixed Issues in Schema Registry.....	62
Fixed Issues in Apache Solr.....	62
Fixed Issues in Spark.....	63
Fixed Issues in Apache Sqoop.....	63
Fixed Issues in Streams Messaging Manager.....	64
Fixed Issues in Streams Replication Manager.....	64
Fixed Issues in Apache Tez.....	64
Fixed Issues in Apache YARN and YARN Queue Manager.....	65
Fixed Issues in Zeppelin.....	67
Fixed Issues in Apache ZooKeeper.....	67
Fixed Issues In Cloudera Runtime 7.2.17.100.....	67
Fixed Issues In Cloudera Runtime 7.2.17.200.....	68
Fixed Issues In Cloudera Runtime 7.2.17.300.....	68
Fixed Issues In Cloudera Runtime 7.2.17.400.....	69
Known Issues In Cloudera Runtime 7.2.17.....	69
Known Issues in Apache Atlas.....	69
Known Issues in Apache Avro.....	74
Known Issues in Cloud Connectors.....	74
Known issues in Cruise Control.....	75
Known Issues in Data Analytics Studio.....	76
Known Issues in Apache HBase.....	77
Known Issues in HDFS.....	77
Known Issues in Apache Hive.....	78
Known Issues in Hue.....	79
Known Issues Iceberg.....	82
Known Issues in Apache Impala.....	83
Known Issues in Apache Kafka.....	87
Known Issues in Apache Knox.....	89
Known Issues in Apache Kudu.....	90
Known Issues in Apache Oozie.....	90
Known Issues in Apache Phoenix.....	91
Known Issues in Apache Ranger.....	91
Known Issues in Schema Registry.....	91
Known Issues in Apache Solr.....	93
Known Issues in Apache Spark.....	98
Known Issues in Apache Spark3.....	99
Known Issues for Apache Sqoop.....	99
Known issues in Streams Messaging Manager.....	99

Known Issues in Streams Replication Manager.....	100
Known Issues in MapReduce, Apache Hadoop YARN, and YARN Queue Manager.....	102
Known Issues in Apache Zeppelin.....	104
Known Issues in Apache ZooKeeper.....	104
Behavioral Changes In Cloudera Runtime 7.2.17.....	104
Behavioral Changes in Apache Kafka.....	105
Behavioral changes in Apache Hive.....	105
Behavioral changes in Apache Ranger.....	105
Behavioral Changes in Apache Solr.....	106
Deprecation Notices In Cloudera Runtime 7.2.17.....	106
Deprecation Notices for Apache Kafka.....	106
Deprecation Notices for Spark 2.....	107
Fixed Common Vulnerabilities and Exposures 7.2.17.....	107

Overview

You can review the Release Notes of Cloudera Runtime 7.2.17 for release-specific information related to new features and improvements, bug fixes, deprecated features and components, known issues, and changed features that can affect product behavior.

Cloudera Runtime Component Versions

You must be familiar with the versions of all the components in the Cloudera Runtime 7.2.17 distribution to ensure compatibility of these components with other applications. You must also be aware of the available Technical Preview components and use them only in a testing environment.

Apache Components

Component	Version
Apache Arrow	0.11.1.7.2.17.0-334
Apache Atlas	2.2.0.7.2.17.0-334
Apache Calcite	1.21.0.7.2.17.0-334
Apache Avro	1.8.2.7.2.17.0-334
Apache Flink	1.16.1.1.10.0.0
Apache Hadoop (Includes YARN and HDFS)	3.1.1.7.2.17.0-334
Apache HBase	2.4.6.7.2.17.0-334
Apache Hive	3.1.3000.7.2.17.0-334
Apache Iceberg	1.1
Apache Impala	4.0.0.7.2.17.0-334
Apache Kafka	3.4.0.7.2.17.0-334
Apache Knox	1.3.0.7.2.17.0-334
Apache Kudu	1.15.0.7.2.17.0-334
Apache Livy	0.7.2.7.2.17.0-334
Apache MapReduce	3.1.1.7.2.17.0-334
Apache NiFi	1.21.0.2.2.7.0
Apache NiFi Registry	1.21.0.2.2.7.0
Apache Oozie	5.1.0.7.2.17.0-334
Apache ORC	1.5.1.7.2.17.0-334
Apache Parquet	1.10.99.7.2.17.0-334
Apache Phoenix	5.1.1.7.2.17.0-334
Apache Ranger	2.3.0.7.2.17.0-334
Apache Solr	8.4.1.7.2.17.0-334
Apache Spark	2.4.8.7.2.17.0-334
Apache Spark 3	3.3.2.7.2.17.0-334
Apache Sqoop	1.4.7.7.2.17.0-334
Apache Tez	0.9.1.7.2.17.0-334

Component	Version
Apache Zeppelin	0.8.2.7.2.17.0-334
Apache ZooKeeper	3.5.5.7.2.17.0-334

Other Components

Component	Version
Cruise Control	2.5.116.7.2.17.0-334
Data Analytics Studio	1.4.2.7.2.17.0-334
GCS Connector	2.1.2.7.2.17.0-334
HBase Indexer	1.5.0.7.2.17.0-334
Hive Solr Connector	4.0.0.7.2.17.0-334
Hue	4.5.0.7.2.17.0-334
Search	1.0.0.7.2.17.0-334
Schema Registry	0.10.0.7.2.17.0-334
Spark Solr Connector	3.9.0.7.2.17.0-334
Streams Messaging Manager	2.3.0.7.2.17.0-334
Streams Replication Manager	1.1.0.7.2.17.0-334

Connectors and Encryption Components

Component	Version
HBase connectors	1.0.0.7.2.17.0-334
Hive Meta Store (HMS)	1.0.0.7.2.17.0-334
Hive on Tez	1.0.0.7.2.17.0-334
Hive Warehouse Connector	1.0.0.7.2.17.0-334
Spark Atlas Connector	0.1.0.7.2.17.0-334
Spark Schema Registry	1.1.0.7.2.17.0-334

Using the Cloudera Runtime Maven repository 7.2.17

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

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



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

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

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

```

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

```

Runtime 7.2.17.0-334

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

Project	groupId	artifactId	version
Atlas	org.apache.atlas	atlas-authorization	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-aws-s3-bridge	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-azure-adls-bridge	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-classification-updater	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-client-common	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-client-v1	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-client-v2	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-common	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-distro	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-docs	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-graphdb-api	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-graphdb-common	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-graphdb-janus	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-hdfs-bridge	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-index-repair-tool	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-intg	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-janusgraph-hbase2	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-notification	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-plugin-classloader	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-repository	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-server-api	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	atlas-testtools	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	hbase-bridge	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	hbase-bridge-shim	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	hbase-testing-util	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	hdfs-model	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	hive-bridge	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	hive-bridge-shim	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	impala-bridge	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	impala-bridge-shim	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	impala-hook-api	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	kafka-bridge	2.1.0.7.2.17.0-334

Project	groupId	artifactId	version
Atlas	org.apache.atlas	kafka-bridge-shim	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	navigator-to-atlas	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	sample-app	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	sqoop-bridge	2.1.0.7.2.17.0-334
Atlas	org.apache.atlas	sqoop-bridge-shim	2.1.0.7.2.17.0-334
Avro	org.apache.avro	avro	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-compiler	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-ipc	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-mapred	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-maven-plugin	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-protobuf	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-service-archetype	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-thrift	1.8.2.7.2.17.0-334
Avro	org.apache.avro	avro-tools	1.8.2.7.2.17.0-334
Avro	org.apache.avro	trevni-avro	1.8.2.7.2.17.0-334
Avro	org.apache.avro	trevni-core	1.8.2.7.2.17.0-334
Calcite	org.apache.calcite	calcite-babel	1.21.0.7.2.17.0-334
Calcite	org.apache.calcite	calcite-core	1.21.0.7.2.17.0-334
Calcite	org.apache.calcite	calcite-druid	1.21.0.7.2.17.0-334
Calcite	org.apache.calcite	calcite-kafka	1.21.0.7.2.17.0-334
Calcite	org.apache.calcite	calcite-linq4j	1.21.0.7.2.17.0-334
Calcite	org.apache.calcite	calcite-server	1.21.0.7.2.17.0-334
Calcite	org.apache.calcite	calcite-ubenchmark	1.21.0.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-aliyun	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-annotations	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-archive-logs	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-archives	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-assemblies	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-auth	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-aws	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-azure	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-azure-datalake	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-benchmark	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-build-tools	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-client	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-client-api	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-client-integration-tests	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-client-minicluster	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-client-runtime	3.1.1.7.2.17.0-334

Project	groupId	artifactId	version
Hadoop	org.apache.hadoop	hadoop-cloud-storage	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-common	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-datajoin	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-distcp	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-extras	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-fs2img	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-gridmix	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-hdfs	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-hdfs-client	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-hdfs-httpfs	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-hdfs-native-client	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-hdfs-nfs	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-hdfs-rbf	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-kafka	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-kms	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-app	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-common	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-core	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-hs	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-hs-plugins	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-jobclient	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-nativetask	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-shuffle	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-uploader	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-mapreduce-examples	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-maven-plugins	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-minicluster	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-minikdc	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-nfs	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-openstack	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-resourceestimator	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-rumen	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-sls	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-streaming	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-api	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-applications-distributedshell	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-applications-unmanaged-am-launcher	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-client	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-common	3.1.1.7.2.17.0-334

Project	groupId	artifactId	version
Hadoop	org.apache.hadoop	hadoop-yarn-registry	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-applicationhistoryservice	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-common	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-nodemanager	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-resourcemanager	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-router	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-sharedcachemanager	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-tests	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-timeline-pluginstorage	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-client	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-common	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-server-2	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-tests	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-server-web-proxy	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-services-api	3.1.1.7.2.17.0-334
Hadoop	org.apache.hadoop	hadoop-yarn-services-core	3.1.1.7.2.17.0-334
HBase	org.apache.hbase	hbase-annotations	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-asyncfs	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-checkstyle	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-client	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-client-project	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-common	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-endpoint	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-examples	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-external-blockcache	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-hadoop-compat	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-hadoop2-compat	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-hbtop	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-http	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-it	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-logging	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-mapreduce	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-metrics	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-metrics-api	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-procedure	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-protocol	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-protocol-shaded	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-replication	2.4.6.7.2.17.0-334

Project	groupId	artifactId	version
HBase	org.apache.hbase	hbase-resource-bundle	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-rest	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-rsgroup	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-server	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-shaded-client	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-shaded-client-byo-hadoop	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-shaded-client-project	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-shaded-mapreduce	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-shaded-testing-util	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-shaded-testing-util-tester	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-shell	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-testing-util	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-thrift	2.4.6.7.2.17.0-334
HBase	org.apache.hbase	hbase-zookeeper	2.4.6.7.2.17.0-334
Hive	org.apache.hive	catalogd-unit	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-beeline	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-blobstore	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-classification	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-cli	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-common	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-contrib	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-exec	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-hbase-handler	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-hcatalog-it-unit	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-hpsql	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-iceberg-catalog	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-iceberg-handler	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-iceberg-shading	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-impala	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-custom-serde	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-impala	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-minikdc	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-qfile-kudu	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-test-serde	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-unit	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-unit-hadoop2	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-it-util	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-jdbc	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-jdbc-handler	3.1.3000.7.2.17.0-334

Project	groupId	artifactId	version
Hive	org.apache.hive	hive-jmh	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-kudu-handler	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-llap-client	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-llap-common	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-llap-ext-client	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-llap-server	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-llap-tez	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-metastore	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-parser	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-pre-upgrade	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-serde	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-service	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-service-rpc	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-shims	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-standalone-metastore	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-storage-api	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-streaming	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-testutils	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-udf	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	hive-vector-code-gen	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	kafka-handler	3.1.3000.7.2.17.0-334
Hive	org.apache.hive	patched-iceberg-api	patched-1.1.0.7.2.17.0-334-3.1.3000.
Hive	org.apache.hive	patched-iceberg-core	patched-1.1.0.7.2.17.0-334-3.1.3000.
Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector-spark3_2.12	1.0.0.7.2.17.0-334
Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector_2.11	1.0.0.7.2.17.0-334
Kafka	org.apache.kafka	ci	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-api	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-basic-auth-extension	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-cloudera-authorization-extension	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-cloudera-common	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-cloudera-secret-storage	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-cloudera-security-policies	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-file	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-json	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-mirror	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-mirror-client	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	connect-runtime	3.4.0.7.2.17.0-334

Project	groupId	artifactId	version
Kafka	org.apache.kafka	connect-transforms	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	generator	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-clients	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-cloudera-metrics-reporter_2.12	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-cloudera-metrics-reporter_2.13	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-cloudera-plugins	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-examples	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-group-coordinator	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-log4j-appender	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-metadata	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-raft	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-server-common	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-shell	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-storage	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-storage-api	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-examples	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-scala_2.12	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-scala_2.13	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-test-utils	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0100	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0101	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0102	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0110	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-10	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-11	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-21	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-22	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-23	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-24	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-25	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-26	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-27	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-28	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-30	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-31	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-32	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-33	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka-tools	3.4.0.7.2.17.0-334

Project	groupId	artifactId	version
Kafka	org.apache.kafka	kafka_2.12	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	kafka_2.13	3.4.0.7.2.17.0-334
Kafka	org.apache.kafka	trogdor	3.4.0.7.2.17.0-334
Knox	org.apache.knox	gateway-adapter	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-admin-ui	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-applications	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-cloud-bindings	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-demo-ldap	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-demo-ldap-launcher	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-discovery-ambari	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-discovery-cm	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-docker	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-i18n	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-i18n-logging-log4j	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-i18n-logging-slf4j	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-performance-test	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-ha	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-identity-assertion-common	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-identity-assertion-concat	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-identity-assertion-hadoop-groups	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-identity-assertion-no-does	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-identity-assertion-pseudo	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-identity-assertion-regex	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-identity-assertion-switchcase	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-jersey	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-rewrite	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-rewrite-common	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-rewrite-func-hostmap-static	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-rewrite-func-inbound-query-param	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-rewrite-func-service-registry	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-rewrite-step-encrypt-uri	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-rewrite-step-secure-query	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-authc-anon	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-authz-acls	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-authz-composite	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-clientcert	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-hadoopauth	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-jwt	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-pac4j	1.3.0.7.2.17.0-334

Project	groupId	artifactId	version
Knox	org.apache.knox	gateway-provider-security-preauth	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-shiro	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-provider-security-webappsec	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-release	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-server	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-server-launcher	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-server-xforwarded-filter	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-admin	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-as	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-auth	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-definitions	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-hashicorp-vault	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-hbase	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-health	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-hive	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-idbroker	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-impala	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-jkg	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-knoxsso	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-knoxsout	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-knoxtoken	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-livy	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-metadata	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-nifi	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-nifi-registry	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-remoteconfig	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-rm	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-session	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-storm	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-test	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-tgs	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-vault	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-service-webhdfs	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-shell	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-shell-launcher	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-shell-release	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-shell-samples	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-spi	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-spi-common	1.3.0.7.2.17.0-334

Project	groupId	artifactId	version
Knox	org.apache.knox	gateway-test	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-test-idbroker	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-test-release-utils	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-test-utils	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-topology-hadoop-xml	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-topology-simple	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-util-common	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-util-configinjector	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-util-launcher	1.3.0.7.2.17.0-334
Knox	org.apache.knox	gateway-util-urltemplate	1.3.0.7.2.17.0-334
Knox	org.apache.knox	hadoop-examples	1.3.0.7.2.17.0-334
Knox	org.apache.knox	knox-cli-launcher	1.3.0.7.2.17.0-334
Knox	org.apache.knox	knox-homepage-ui	1.3.0.7.2.17.0-334
Knox	org.apache.knox	knox-token-generation-ui	1.3.0.7.2.17.0-334
Knox	org.apache.knox	knox-token-management-ui	1.3.0.7.2.17.0-334
Knox	org.apache.knox	knox-webshell-ui	1.3.0.7.2.17.0-334
Knox	org.apache.knox	webhdfs-kerb-test	1.3.0.7.2.17.0-334
Knox	org.apache.knox	webhdfs-test	1.3.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-backup-tools	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-backup2_2.11	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-backup3_2.12	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-client	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-hive	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-spark2-tools_2.11	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-spark2_2.11	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-spark3-tools_2.12	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-spark3_2.12	1.15.0.7.2.17.0-334
Kudu	org.apache.kudu	kudu-test-utils	1.15.0.7.2.17.0-334
Livy	org.apache.livy	livy-api	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-client-common	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-client-http	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-core_2.12	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-examples	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-integration-test	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-repl_2.11	0.7.2.7.2.17.0-334
Livy	org.apache.livy	livy-repl_2.12	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-rsc	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-scala-api_2.11	0.7.2.7.2.17.0-334
Livy	org.apache.livy	livy-scala-api_2.12	0.7.23000.7.2.17.0-334

Project	groupId	artifactId	version
Livy	org.apache.livy	livy-server	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-test-lib	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-thriftserver	0.7.23000.7.2.17.0-334
Livy	org.apache.livy	livy-thriftserver-session	0.7.23000.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-common	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-icu	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-kuromoji	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-morfologik	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-nori	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-openslp	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-phonetic	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-smartcn	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-analyzers-stempel	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-backward-codecs	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-benchmark	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-classification	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-codecs	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-core	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-demo	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-expressions	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-facet	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-grouping	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-highlighter	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-join	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-memory	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-misc	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-monitor	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-queries	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-queryparser	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-replicator	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-sandbox	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-spatial	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-spatial-extras	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-spatial3d	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-suggest	8.4.1.7.2.17.0-334
Lucene	org.apache.lucene	lucene-test-framework	8.4.1.7.2.17.0-334
Oozie	org.apache.oozie	oozie-client	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-core	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-distro	5.1.0.7.2.17.0-334

Project	groupId	artifactId	version
Oozie	org.apache.oozie	oozie-examples	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-fluent-job-api	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-fluent-job-client	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-server	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-distcp	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-git	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-hcatalog	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-hive	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-hive2	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-oozie	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-spark	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-sqoop	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-sharelib-streaming	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-tools	5.1.0.7.2.17.0-334
Oozie	org.apache.oozie	oozie-zookeeper-security-tests	5.1.0.7.2.17.0-334
ORC	org.apache.orc	orc-core	1.5.1.7.2.17.0-334
ORC	org.apache.orc	orc-examples	1.5.1.7.2.17.0-334
ORC	org.apache.orc	orc-mapreduce	1.5.1.7.2.17.0-334
ORC	org.apache.orc	orc-shims	1.5.1.7.2.17.0-334
ORC	org.apache.orc	orc-tools	1.5.1.7.2.17.0-334
Parquet	org.apache.parquet	parquet-avro	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-cascading	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-cascading3	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-column	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-common	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-encoding	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-format-structures	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-generator	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-hadoop	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-hadoop-bundle	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-jackson	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-pig	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-pig-bundle	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-protobuf	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-scala_2.10	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-thrift	1.10.99.7.2.17.0-334
Parquet	org.apache.parquet	parquet-tools	1.10.99.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-client-embedded-hbase-2.4	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-client-hbase-2.4	5.1.1.7.2.17.0-334

Project	groupId	artifactId	version
Phoenix	org.apache.phoenix	phoenix-connectors-phoenix5-compat	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-core	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.1.6	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.2.5	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.3.0	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.4.0	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.4.1	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-pherf	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-queryserver	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-queryserver-client	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-queryserver-it	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-queryserver-load-balancer	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-queryserver-orchestrator	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-server-hbase-2.4	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix-tracing-webapp	5.1.1.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix5-hive	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix5-hive-shaded	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix5-spark	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix5-spark-shaded	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix5-spark3	6.0.0.7.2.17.0-334
Phoenix	org.apache.phoenix	phoenix5-spark3-shaded	6.0.0.7.2.17.0-334
Ranger	org.apache.ranger	conditions-enrichers	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	credentialbuilder	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	jisql	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ldapconfigcheck	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-adls-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-atlas-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-atlas-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-authn	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-common-ha	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-distro	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-examples-distro	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-gs-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-hbase-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-hbase-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-hdfs-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-hdfs-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-hive-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-hive-plugin-shim	2.3.0.7.2.17.0-334

Project	groupId	artifactId	version
Ranger	org.apache.ranger	ranger-intg	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kafka-connect-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kafka-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kafka-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kms	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kms-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kms-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-knox-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-knox-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kudu-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kylin-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-kylin-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-metrics	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-nifi-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-nifi-registry-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-ozone-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-ozone-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-plugin-classloader	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-plugins-audit	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-plugins-common	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-plugins-cred	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-policymigration	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-adls	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-chained-plugins	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-hook-abfs	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-hook-s3	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-intg	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-processor	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-s3	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-raz-s3-lib	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-rms-common	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-rms-hive	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-rms-plugins-common	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-rms-webapp	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-s3-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-sampleapp-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-schema-registry-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-solr-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-solr-plugin-shim	2.3.0.7.2.17.0-334

Project	groupId	artifactId	version
Ranger	org.apache.ranger	ranger-sqoop-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-sqoop-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-storm-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-storm-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-tagsync	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-tools	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-util	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-yarn-plugin	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ranger-yarn-plugin-shim	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	sample-client	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	sampleapp	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	shaded-raz-hook-abfs	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	shaded-raz-hook-s3	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	ugsync-util	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	unixauthclient	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	unixauthservice	2.3.0.7.2.17.0-334
Ranger	org.apache.ranger	unixusersync	2.3.0.7.2.17.0-334
Solr	org.apache.solr	solr-analysis-extras	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-analytics	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-cell	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-clustering	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-core	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-dataimporthandler	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-dataimporthandler-extras	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-jaegertracer-configurator	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-langid	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-ltr	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-prometheus-exporter	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-security-util	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-solrj	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-test-framework	8.4.1.7.2.17.0-334
Solr	org.apache.solr	solr-velocity	8.4.1.7.2.17.0-334
Spark	org.apache.spark	spark-avro_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-avro_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-catalyst_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-catalyst_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-core_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-core_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-graphx_2.11	2.4.8.7.2.17.0-334

Project	groupId	artifactId	version
Spark	org.apache.spark	spark-graphx_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-hadoop-cloud_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-hadoop-cloud_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-hive_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-hive_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-kubernetes_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-kubernetes_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-kvstore_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-kvstore_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-launcher_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-launcher_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-mllib-local_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-mllib-local_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-mllib_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-mllib_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-network-common_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-network-common_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-network-shuffle_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-network-shuffle_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-network-yarn_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-network-yarn_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-repl_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-repl_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-shaded-raz	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-sketch_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-sketch_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-sql-kafka-0-10_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-sql-kafka-0-10_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-sql_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-sql_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-streaming-kafka-0-10_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-streaming-kafka-0-10_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-streaming_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-streaming_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-tags_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-tags_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-token-provider-kafka-0-10_2.11	2.4.8.7.2.17.0-334

Project	groupId	artifactId	version
Spark	org.apache.spark	spark-token-provider-kafka-0-10_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-unsafe_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-unsafe_2.12	3.3.2.7.2.17.0-334
Spark	org.apache.spark	spark-yarn_2.11	2.4.8.7.2.17.0-334
Spark	org.apache.spark	spark-yarn_2.12	3.3.2.7.2.17.0-334
Sqoop	org.apache.sqoop	sqoop	1.4.7.7.2.17.0-334
Sqoop	org.apache.sqoop	sqoop-test	1.4.7.7.2.17.0-334
Tez	org.apache.tez	hadoop-shim	0.9.1.7.2.17.0-334
Tez	org.apache.tez	hadoop-shim-2.8	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-api	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-aux-services	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-common	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-dag	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-examples	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-ext-service-tests	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-history-parser	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-javadoc-tools	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-job-analyzer	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-mapreduce	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-protobuf-history-plugin	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-runtime-internals	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-runtime-library	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-tests	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-yarn-timeline-cache-plugin	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-yarn-timeline-history	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-yarn-timeline-history-with-acls	0.9.1.7.2.17.0-334
Tez	org.apache.tez	tez-yarn-timeline-history-with-fs	0.9.1.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-angular	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-display	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-interpreter	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-jdbc	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-jupyter	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-livy	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-markdown	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-server	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-shaded-raz	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-shell	0.8.2.7.2.17.0-334
Zeppelin	org.apache.zeppelin	zeppelin-zengine	0.8.2.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper	3.5.5.7.2.17.0-334

Project	groupId	artifactId	version
ZooKeeper	org.apache.zookeeper	zookeeper-client-c	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-loggraph	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-rest	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-docs	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-jute	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-election	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-lock	3.5.5.7.2.17.0-334
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-queue	3.5.5.7.2.17.0-334

What's New In Cloudera Runtime 7.2.17

You must be aware of the additional functionalities and improvements to features of components in Cloudera Runtime 7.2.17. Learn how the new features and improvements benefit you.

What's New in Apache Atlas

Learn about the new features of Apache Atlas in Cloudera Runtime 7.2.17.

Improved search capabilities

Atlas supports searching for entities using more options..You can download search results for both basic and advanced search flavors.

See [Ability to download search results](#) for more information.

Iceberg support for Atlas (Technical Preview)

Atlas integration with Iceberg helps you identify the Iceberg tables to scan data and provide lineage support.

Iceberg for Atlas feature is available within Cloudera Runtime 7.2.17 but is not ready for production deployment. Cloudera encourages you to explore this technical preview feature in non-production environments and provide feedback on your experiences through the Cloudera Community Forums.

See [Iceberg for Atlas](#) for more information.

AttributeName validation in parent and child TypeDef

Atlas service validates the attribute names of the entity types for those attributes having identical names as their parents' attributes.

See [Validating the AttributeName in parent and child TypeDef](#) for more information.

Improved dynamic indexy recovery

The JanusGraph database transaction might fail in certain scenarios and this failure can be handled dynamically using a specific configuration.

See [Dynamic handling of failure](#) for more information.

FISMA compliance support for Atlas

FISMA requires the use of FIPS 140 validated crypto modules for encrypted storage (not just FIPS compliant but a valid FIPS certificate must be provided). In order to achieve this in Atlas, the data stored in HBASE must be setup to be encrypted when the tables are created by Atlas. The user needs to enable encryption for the EBS volumes used for HBASE storage. Users looking for encrypted HBASE storage must take required steps to ensure that the HBASE storage they would be using is encrypted.

Log4j updates

Log4j migrated from 1.x to 2.x. All log4j 1.x dependencies have been removed to address CVEs and EOL

What's New in Cruise Control

Learn about the new features of Cruise Control in Cloudera Runtime 7.2.17.

Rebasing Cruise Control to 2.5.116

Cruise Control in Cloudera Runtime is rebased to the 2.5.116 version. For more information about the fixes and features in Cruise Control 2.5.116, see the [Cruise Control Rebase Summary](#).

New endpoint for Cruise Control

The GET/kafkacruisecontrol/permissions endpoint is added to Cruise Control that lists the level permissions of the current user. In case authentication is not configured for a user, the GET call returns Unable to retrieve privilege information for an unsecure connection message.

Cruise Control Rebase Summary

In Cloudera Runtime 7.2.17, Cruise Control is rebased to the 2.5.116 version. Other than the added new feature, several issues are fixed and several features are enhanced to have a better performance when using Cruise Control.

Table 1: Fixed Issues

PR-1758 Fix request log to write to configured CruiseControlPublicAccessLogger
PR-1847 Fix minor logging issues
PR-1875 Fix the BrokerFailureDetector doesn't work issue
PR-1901 Fix failedBrokers.txt permissions
PR-1939 Address issues with Vertx based Swagger UI
PR-1949 Fix the response code to 202 for in-progress request

Table 2: Version Update

PR-1810 Bump vulnerable transitive jackson-databind dependency
PR-1811 Bump jackson-databind 2.12.6.1 to resolve CVE-2020-36518
PR-1815 Bump swagger-parser for CVE-2020-36518
PR-1855 Upgrade Netty and Jetty versions for CVE fixes
PR-1861 Downgrade netty and jetty dependencies to a safer version
PR-1902 Bump fastxml dependencies to fix a High CVE score for the org.yaml:snakeyaml package
PR-1926 chore: bump fastxml dependencies to fix a High CVE score
PR-1928 Bumped dependency swagger-parser-v3 version to latest 2.1.3
PR-1937 Bump scala version to 2.13.10 (CVE-2022-36944)

Table 3: Goal Improvements

PR-1809 Add BrokerSetAwareGoal
PR-1857 Add more details to the goals config description
PR-1919 Add fixability metrics to GoalOptimizer

Table 4: Feature Maintenance

PR-1789 Update BrokerFailureDetector to use AdminClient clusters	PR-1897 Add more logging to SlowBrokerFinder
PR-1803 Cleanups for running without ZooKeeper	PR-1895 Rename the getGoalsByPriority to getDefaultGoalsByPriority
PR-1825 Remove execution tasks that has in-movement partitions from re-execution candidate	PR-1889 Display the brokerSet id in kafka_cluster_state endpoint
PR-1830 Make the retrieval of the desired replication factor of sample store topics more robust	PR-1910 Future-proof for Kafka 3.3 for KafkaYammerMetrics class name changes
PR-1831 Allow fraction CPU core capacity values	PR-1914 Fix leader replica cpu util when leader egress is zero
PR-1839 Update the numCore type to double in BasicStats	PR-1921 Add Vertx.io based API with swagger UI
PR-1841 Update NumCore type in brokerStats.yaml	PR-1941 Log broker-set resolution error only when BrokerSetAwareGoal is in default goal list
PR-1848 Enable users to get remote-storage-enabled of topics in kafka_cluster_state	PR-1967 Add metrics to reflect the partition movement speed
PR-1867 Make Prometheus broker cpu metric query configurable	PR-1968 Implement concurrency adjuster for each individual broker
PR-1879 Handle the inconsistency between clusterModel and kafka metadata during update topic configuration	PR-1977 Add brokerSet to JSON API response
PR-1881 Mute KafkaCruiseControlConfig logs during anomaly detection	PR-1980 Dynamically adjust the executionProgressCheckIntervalMs based on execution status
PR-1884 Make sure the provision response is always RIGHT_SIZED if it is not OVER_PROVISIONED	PR-1983 Add initialized state to concurrency adjuster and avoid null for the metric value
PR-1891 Make the state output easier to read	PR-1985 Remove initialized check when get concurrency summary

What's new in Data Analytics Studio

Learn about what is new in Data Analytics Studio (DAS) in Cloudera Runtime 7.2.17.

DAS has been deprecated

DAS has been deprecated in Cloudera Data Hub (Cloudera Runtime 7.2.17 release) and will be removed from the CDP stack in future releases. Cloudera encourages you to use Hue to run Hive LLAP workloads. You can submit queries from DAS, but cannot view query history without enabling the DAS Event Processor. For more information, see [Enabling the DAS Event Processor](#).

What's New in Apache HBase

Learn about the new features of HBase in Cloudera Runtime 7.2.17.

The initial corePoolSize in HBase is configurable for ChoreService

The `hbase.choreservice.initial.pool.size` configuration property is added to HBase to set the initial number of threads for the ChoreService. The default value is 1.

The related Apache HBase JIRA: [HBASE-27565](#).

Disable sorting the directories by size in HBase CleanerChore service

The `hbase.cleaner.directory.sorting` configuration property is added to HBase so that the CleanerChore service can be disabled and sort the sub directories by the consumed space, and start the cleaning with the largest sub directory. The property is enabled by default.

The related Apache HBase JIRA: [HBASE-27506](#).

Port FileWatcher from Zookeeper to detect keystore or truststore changes in TLS connections automatically

HBase detects the changes in the filesystem to refresh the keystore or truststore automatically during runtime. Now you do not need to restart the HBase instances manually when certificates are refreshed.

The related Apache HBase JIRA: [HBASE-27347](#).

Automatically detect the keystore or truststore file types using file extension

HBase detects the keystore or truststore file types using the file extension if it is not explicitly specified in the configuration. If the file type cannot be detected, JKS is used by default.

The related Apache HBase JIRA: [HBASE-27346](#).

Add mutual authentication support to TLS

By default, when TLS is enabled, mutual authentication of certificates is enabled. This means, during handshake, the client authenticates the server's certificate (as usual) and also the server authenticates the client's certificate. Additionally, each side validates that the hostname presented by the certificate matches the address of the connection.

- `hbase.server.netty.tls.client.auth.mode`: Default value is NEED. Possible values are, NEED, WANT, NONE.
- `hbase.server.netty.tls.verify.client.hostname`: Default value is true.
- `hbase.client.netty.tls.verify.server.hostname`: Default value is true.

Additionally, during hostname verification, if required a fallback on reverse lookup is supported. The reverse lookup can be disabled using `hbase.rpc.tls.host-verification.reverse-dns.enabled` property. The default value is true.

The related Apache HBase JIRA: [HBASE-27280](#).

What's New in Apache Hive

Learn about the new features of Hive in Cloudera Runtime 7.2.17.

Hive Metastore dynamic leader election

You can enable dynamic leader election for Hive Metastore (HMS) to avoid running the same tasks across all Hive Metastore (HMS) instances. HMS performs housekeeping tasks, such as execution of compaction tasks, auto-discovering partitions for external tables, generation of compaction tasks, and so on. When there are multiple HMS instances, it is essential to have a single leader HMS elected to avoid running the same tasks across all the instances. The elected leader then performs the housekeeping tasks.

This feature uses Hive locks to dynamically elect a leader. When a HMS instance owns a lock, it is elected as the leader and performs the housekeeping tasks. For more information, see [Hive Metastore leader election](#).

What's New in Hue

Learn about the new features of Hue in Cloudera Runtime 7.2.17.

Performance and security enhancements in Hue

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

- Python libraries such as django-auth-ldap, django-axes,.djangorestframework-simplejwt, Mako, Markdown, python-ldap, django-babel, django-mako, django-cors-headers,.djangorestframework, eventlet, sqlparse, and so on have been upgraded from Python 2.7 to Python 3.8.
- The Django server has been upgraded from version 1.11.29 to 3.2.16.
- Hue now uses Gunicorn as a front-end server. Previously, Hue used the CherryPy server.

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

Ability to set the permitted file size for upload using Hue File Browser

You can set the permitted size of a file that your users can upload using the Hue File Browser by setting the following parameter in the Hue Advanced Configuration in Cloudera Manager:

```
[filebrowser]
max_file_size_upload_limit=[**FILE-SIZE-IN-BYTES**] \\default is -1 (no
limit)
```

Ability to access S3 buckets from Hue with RAZ is GA

Granting fine-grained access to the per-user home directories in Amazon S3 and accessing them from the S3 File Browser in Hue using RAZ is GA. See [Enabling S3 File Browser for Hue with RAZ in DataHub](#).

Ability to access ADLS Gen2 containers from Hue with RAZ is GA

Granting fine-grained access to the per-user home directories on ADLS Gen2 containers and accessing them from the ABFS File Browser in Hue using RAZ is GA. See [Enabling ABFS File Browser in Hue with RAZ in DataHub](#).

Increased the download limit on the Solr dashboard

Earlier, you could download only 1000 records from the Solr Search dashboard. Hue now supports downloading up to 15000 records. You can configure the download limit using the following Advanced Configuration Snippet:

```
[search]
download_limit=[**DOWNLOAD-LIMIT**]
```

What's new in Apache Iceberg

Learn about the new features of Iceberg in Cloudera Runtime 7.2.17.

This release introduces the general availability of ACID transactions with Iceberg v2 tables from Hive in CDP Runtime 7.2.17. You can run Apache Iceberg ACID transactions within some of the key data services in the Cloudera Data Platform (CDP) public cloud. From Hive or Impala, you use Apache Iceberg features in Data Hub, which include time travel, create table as select, and schema and partition evolution. For more information about using Iceberg, see "[Using Iceberg](#)".

What's New in Apache Impala

Learn about the new features of Impala in Cloudera Runtime 7.2.17.

View query timeline in Impala WebUI

For a detailed report on how a query was executed and to understand the detailed query performance characteristics, use the built-in web server's UI and look at the [Gantt chart](#).

Ability to create a non-unique primary key for Kudu

Impala now supports [creating a Kudu table with a non-unique primary key](#). When creating a Kudu table, specifying PRIMARY KEY is optional now. If there is no primary key attribute specified, the partition key columns could be promoted as non-unique primary keys if those columns are the beginning columns of the table.

Improvement in Ozone file handle caching

In this release, [Ozone file handle caching](#) is enabled by default. This file handle caching improves TPC-DS geomean query time by 10%.

Support Read from Ozone data with erasure coding

Impala now supports [reading from Ozone data stored with Erasure Coding \(EC\)](#). The Ozone EC feature provides data durability and fault tolerance with reduced storage space and ensures data durability similar to the Ratis THREE replication approach. EC can be considered as an alternative to replication.

Spill to HDFS

Impala occasionally needs to use persistent storage for writing intermediate files during large sorts, joins, aggregations, or analytical function operations. If your workload results in large volumes of intermediate written data, it is recommended to configure the heavy spilling queries to use a remote storage location rather than the local one. The advantage of using [remote storage for scratch space](#) is that it is elastic and can handle large amounts of spilling.

New bytes-read-encrypted metric

When Impala executes any query on an Ozone encrypted file, the query PROFILE captures the runtime details of the execution, including the total number of [encrypted bytes read from Ozone](#) by the query.

Spill to Ozone

You can now use [Ozone as a scratch space](#) for writing intermediate files during large sorts, joins, aggregations, or analytic function operations.

New metrics for EC reads

The [Query Details page](#) contains the low-level details of how a SQL query is processed through Impala. You can now use the Query details page to view the erasure-coded bytes read metrics.

Ability to configure impala-shell client retry attempts

From this release, you can [configure the maximum number of attempts](#) impala-shell can make using an impala-shell option.

Ability to CREATE/ALTER VIEW SET/UNSET TBLPROPERTIES

Before this release, altering only the VIEW definition, VIEW name, and owner was supported. Impala now supports [altering the table properties](#) of a VIEW by using the set tblproperties clause.

Flags related to use_local_catalog

When use_local_catalog is enabled or set to True on the impalad coordinators a new [list of flags](#) configures various parameters as described here. It is not recommended to change the default values on these flags.

Synchronization between Impala clusters

When a LOAD statement is run from Impala in a cluster, an [INSERT event](#) is generated. This INSERT event generated notifies other Impala clusters or any other systems that consume HMS events (e.g. Hive Replication) about the changes of LOAD DATA. The other Impala clusters will refresh the table or partitions based on the event.

Support for all complex types in a SELECT * query

A SELECT * statement did not expand to complex types to be compatible with earlier versions of Impala that did not support complex types in the result set. In the older versions of Impala, queries using SELECT * skip complex types by default and only expanded to primitive types even when the table contained complex-typed columns. This release adds a new query option EXPAND_COMPLEX_TYPES to include [complex types](#) in the SELECT * list.

Allow map type in the SELECT list

You can now use [a SELECT statement to run queries](#) on the keys and values of maps. However, you cannot have mixed complex types in the select list such as collections (arrays or maps) in structs or structs in collections. Also, sorting is not supported if the select list contains collection columns.

Push down date literals to Kudu scanner

Impala now allows [creating and pushing down Kudu predicates](#) from the DATE type.

Example:

```
select * from functional_kudu.date_tbl where date_col = DATE "1970-01-01";

PLAN-ROOT SINK
|
00:SCAN KUDU [functional_kudu.date_tbl]
kudu predicates: date_col = DATE '1970-01-01'
row-size=12B cardinality=1
---- DISTRIBUTEDPLAN
PLAN-ROOT SINK
|
01:EXCHANGE [UNPARTITIONED]
|
00:SCAN KUDU [functional_kudu.date_tbl]
kudu predicates: date_col = DATE '1970-01-01'
row-size=12B cardinality=1
```

UTF-8 mode support

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

Binary support

Impala now supports BINARY columns for all table formats except Kudu. See the [BINARY support topic](#) for more information on using this arbitrary-length byte array data type in CREATE TABLE and SELECT statements.

Increased `data_cache_write_concurrency` default for SSDs

To avoid overwhelming the underlying IO device, the data cache limits concurrent writes to the cache. This is controlled by the `data_cache_write_concurrency` flag and it defaults to 1. If the underlying IO device is identified to be an SSD, the default value of this flag will be increased to allow more concurrent writes to the data cache. This would allow the data cache to warm up faster and stay more up-to-date. When the default value is used, the rotational disks continue to use a default of 1, while non-rotational disks use a default of 8.

Support custom hash schema for Kudu range tables

Impala now includes [CREATE TABLE and ALTER TABLE syntax](#) to allow Kudu custom hash schema. HASH syntax within a partition is similar to the table-level syntax except that HASH clauses must follow the PARTITION clause and commas are not allowed within a partition. You can use the SHOW HASH SCHEMA statement to view the hash schema information for each partition.

Example:

```
CREATE TABLE t1 (id int, c2 int, PRIMARY KEY(id, c2))
PARTITION BY HASH(id) PARTITIONS 3 HASH(c2) PARTITIONS 4
RANGE (c2)
(
  PARTITION 0 <= VALUES < 10
  PARTITION 10 <= VALUES < 20
  HASH(id) PARTITIONS 2 HASH(c2) PARTITIONS 3
  PARTITION 20 <= VALUES < 30
)
STORED AS KUDU;
ALTER TABLE t1 ADD RANGE PARTITION 30 <= VALUES < 40
HASH(id) PARTITIONS 3 HASH(c2) PARTITIONS 4;
```

What's New in Apache Kafka

Learn about the new features of Apache Kafka in Cloudera Runtime 7.2.17.

Rebase on Kafka 3.4.0

Kafka shipped with this version of Cloudera Runtime is based on Apache Kafka 3.4.0. For more information, see the following upstream resources:

Apache Kafka Notable Changes:

- [3.2.0](#)
- [3.3.0 and 3.3.1](#)
- [3.4.0](#)

Apache Kafka Release Notes:

- [3.2.0](#)
- [3.3.0](#)
- [3.3.1](#)
- [3.4.0](#)

Kafka KRaft [TECHNICAL PREVIEW]

Apache Kafka Raft (KRaft) is a consensus protocol used for metadata management that was developed as a replacement for Apache ZooKeeper. Using KRaft for managing Kafka metadata instead of ZooKeeper offers various benefits including a simplified architecture and a reduced operational footprint.

Kafka KRaft in this release of Cloudera Runtime is in technical preview and does not support the following:

- Deployments with multiple log directories. This includes deployments that use JBOD for storage.
- Delegation token based authentication.
- Migrating an already running Kafka service from ZooKeeper to KRaft.
- Atlas Integration.

For a conceptual overview on KRaft, see [Kafka KRaft](#). For more information on how to deploy a Streams Messaging Data Hub cluster that is running KRaft mode, see [Setting up your Streams Messaging cluster](#).

SMT plugins for binary conversion

Two Cloudera developed Single Message Transforms (SMT) plugins are added. These are the ConvertToBytes and ConvertFromBytes plugins, which you can use to convert binary data to or from the Kafka Connect internal data format.

For more information, see the following resources:

- [Single Message Transforms](#)
- [ConvertFromBytes](#)
- [ConvertToBytes](#)

EOS for source connectors

Exactly-once semantics (EOS) support is added for Kafka Connect source connectors. For more information, see [Configuring EOS for source connectors](#).

Rolling restart checks provide a high cluster health guarantees by default

The default value of the Cluster Health Guarantee During Rolling Restart property is changed from none to healthy partitions stay healthy. This property defines what type of checks are performed during a Rolling Restart on the restarted broker. Each setting guarantees a different level of cluster health during Rolling Restarts. With the none setting, no checks are performed. This means that in previous versions no guarantees were provided on cluster health by default.

The new default, healthy partitions stay healthy, ensures a high level of guarantees on cluster health. This setting ensures that no partitions go into an under-min-isr state when a broker is stopped. This is achieved by waiting before each broker is stopped so that all other brokers can catch up with all replicas that are in an at-min-isr state. Additionally, the setting ensures that the restarted broker is accepting requests on its service port before restarting the next broker. This setting ignores partitions which are already in an under-min-isr state. For more information, see [Configuring EOS for source connectors](#).

LDAPS SSL configurations are inherited from the Kafka broker

The SSL configurations of LDAP over SSL (LDAPS) are inherited from the Kafka broker. Previously, the JDK default was used. If the JDK default certificate store contains certificates which were used to setup SSL connection to LDAP, it should be imported to the broker stores.

Aliases for Kafka CLI tools

Aliases are added for the kafka-storage.sh, kafka-cluster.sh, and kafka-features.sh command line tools. These tools can now be called globally with kafka-storage, kafka-cluster, and kafka-features.



Important: Not all tools are fully supported and their use is limited. For more information, see [Unsupported command line tools](#).

What's New in Apache Knox

Learn about the new features of Knox customers in Cloudera Runtime 7.2.17:

Performance and Function Improvements

Listed under Fixed Issues for Knox.

Custom Knox Topologies

Custom descriptors can now be deployed to Apache Knox using Cloudera Manager. These descriptors, combined with referenced provider configurations, are transformed into Knox topologies. Using Cloudera Manager means that these descriptors only ever need to be changed in one place to affect all Knox Gateway instances in the cluster. See [Add a custom descriptor to Apache Knox](#).

What's New in Apache Kudu

Learn about the new features of Kudu in Cloudera Runtime 7.2.17.

Auto-leader rebalancing

An experimental feature is added to Kudu that allows it to automatically rebalance tablet leader replicas among tablet servers. The background task can be enabled by setting the `--auto_leader_rebalancing_enabled` flag on the Kudu masters (see, *KUDU-3390*).

Immutable column

Introduced immutable column. It is useful to define such a column which represents a semantically constant entity (see, *KUDU-3353*).

Auto-incrementing column

Introduced auto-incrementing column. These columns are populated on the server side with a monotonically increasing counter. The counter is local to every tablet; for example, each tablet has a separate auto incrementing counter.

Kudu now supports experimental non-unique primary key. When a table with non-unique primary key is created, an auto-incrementing column named `auto_incrementing_id` will be added automatically to the table as the key column. The non-unique key columns and the auto-incrementing column together form the effective primary key (see, *KUDU-1945*). For more details, see [Non-unique primary key index](#).

Kudu JWT support and proxy support

JWT authentication is an alternative to Kerberos authentication, and you can use it in situations where Kerberos authentication is not a viable option but authentication is required nevertheless. For more details, see [Configuring JWT authentication for Kudu](#).

It is now possible to separate the internal and the external traffic in a Kudu cluster while providing the connectivity for Kudu clients running in external networks where the internal traffic is never routed through a proxy's or a loadbalancer's endpoint. Essentially, it allows for the internal traffic (for example, the traffic between tablet servers and masters) to bypass advertised RPC addresses, using alternative addresses for inter-cluster communications. For more details, see [Proxied RPCs in Kudu](#).

Added sanity check to detect wall clock jumps

Added a sanity check to detect strange jumps in wall clock readings. The idea is to rely on the readings from the `CLOCK_MONOTONIC_RAW` clock captured along with the wall clock readings. A jump should manifest itself in a big difference between the wall clock delta and the corresponding `CLOCK_MONOTONIC_RAW` delta. If such a condition is detected, then `HybridClock::NowWithErrorUnlocked()` dumps diagnostic information about clock NTP synchronisation status and returns `Status::ServiceUnavailable()` with appropriate error message.

As a part of this changelist, the following new flags are introduced:

- `--wall_clock_jump_detection`

This is to control the newly introduced sanity check for readings of the wall clock. Acceptable values are auto, enabled, and disabled. It is set to auto by default, which means that the sanity check for timestamps is enabled if the process detects that it is running on a VM in Azure cloud.

- `--wall_clock_jump_threshold_sec`

This is to control the threshold (in seconds) for the difference in deltas of the wall clock's and `CLOCK_MONOTONIC_RAW` clock's readings. It is set to 900 (15 minutes) by default.

Improvements

None.

What's New in Apache Phoenix

Learn about the new features of Phoenix in Cloudera Runtime 7.2.17.

Phoenix is FIPS compliant

Phoenix is now Federal Information Processing Standards (FIPS) compliant. For more information, see [Phoenix is FIPS compliant](#).

Phoenix thin client supports multiple URLs in JDBC connection parameters

When a Phoenix thin client connects to a PQS, you can now specify a list of servers in the JDBC connection string. Client connects to one of those URLs and if the connection is unsuccessful, client attempts failover (connects to a different URL).

What's New in Apache Ranger

No new features were made generally available for Ranger customers in Cloudera Runtime 7.2.17:

Performance and Function Improvements

Listed under Fixed Issues for Ranger.

GCS Fine-grained Access Control (Preview)

You can now register a CDP environment on GCP with RAZ enabled to use fine-grained access policies and audit capabilities available in Apache Ranger. See [GCS Fine-Grained Access Control](#).



Note: You need to contact Cloudera to have this feature enabled.

What's New in Schema Registry

Learn about the new features of Schema Registry in Cloudera Runtime 7.2.17.

KafkaAvroSerializer and KafkaAvroDeserializer improvements

KafkaAvroSerializer and KafkaAvroDeserializer can now handle null values without Avro

The `KafkaAvroSerializer` and `KafkaAvroDeserializer` now support a configuration property called `null.passthrough.enabled`, which is false by default. If enabled, null data is handled as null, and no schema is sent to Schema Registry. This behavior enables client applications to write tombstone

messages into compact topics. The `KafkaAvroDeserializer` also handles null values by returning null without any regards to the schema.

Support deserialization when the topic and schema names don't match

From now on, the `KafkaAvroDeserializer` uses the schema version's ID in the Avro byte stream to access the actual schema and disregards schema names.

Logical types conversion for the `KafkaAvroSerializer` and `KafkaAvroDeserializer`

The `KafkaAvroSerializer` and `KafkaAvroDeserializer` can now properly handle and convert Avro logical types at a record level. This means that if you have a record that has a field with a built-in Avro logical type (for example a `BigDecimal` field with `BYTES` type and decimal logical type), you can now properly serialize the records. After deserialization, a `GenericRecord` is returned, including the typed `BigDecimal` field, instead of a `ByteBuffer`. Logical type conversion can be enabled using the `logical.type.conversion.enabled` property. This property is set to false by default for backward compatibility.

For more information, see the following resources:

- [KafkaAvroDeserializer properties reference](#)
- [KafkaAvroSerializer properties reference](#)

Principal mapping rules can be defined without quotes

The SSL Client Authentication Mapping Rules (`schema.registry.ssl.principal.mapping.rules`) property now accepts rules that are defined without quotes. As a result, when adding multiple rules, you no longer need to enclose each rule in quotes.

Remove modules section from `registry.yaml`

In previous versions, the `registry.yaml` configuration file contained a `modules` section. This section was used to list pluggable modules that extended Schema Registry's functionality. However, modules were never fully supported and have been removed in a previous release. The `modules` section in `registry.yaml` was kept for backwards compatibility. Starting with this version, the `modules` section is removed by default from `registry.yaml`.

What's New in Apache Spark

Learn about the new features of Spark in Cloudera Runtime 7.2.17.

Apache Spark 3 version support

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

See [Running Apache Spark 3 applications](#) and [Data Engineering clusters](#).

What's New in Sqoop

Learn what's new in the Apache Sqoop client in Cloudera Runtime 7.2.17.

To access the latest Sqoop documentation on Cloudera's documentation web site, go to [Sqoop Documentation 1.4.7.7.1.6.0](#).

Discontinued maintenance of direct mode

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

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

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

Sqoop direct mode is disabled by default. However, if you still want to use it, enable it by either setting the `sqoop.enable.deprecated.direct` property globally in Cloudera Manager for Sqoop or by specifying it in the command-line through `-Dsqoop.enable.deprecated.direct=true`.


What's new in Streams Messaging Manager

Learn about the new features of Streams Messaging Manager in Cloudera Runtime 7.2.17.

UI updates

The style of SMM UI is updated. This update includes various changes to the colors, fonts, and overall style of the UI. Additionally, the following functional changes and improvements are made:

Data Explorer

- The modal window that you use to view messages now includes a copy to clipboard button if the message you are viewing is long.
- A  (Refresh) option is added next to the FROM OFFSET field. This option refreshes the partition offset range and fetches the latest messages.

Connector Configuration and Connector Settings pages

- A new option, Add, is added to the **Import a Connector config...** modal. This option enables you to import connector configuration properties without overriding existing properties.
- Property keys can now be filtered based on their group and importance.
- A Reset Filters option is added, this option resets all search filters.
- Three new actions are added that modify the configuration as a whole. The options are Remove all, Reset, and Export. These actions are available in a new Actions drop-down.
- The Import Connector Configuration... option is moved to the Actions drop-down and is renamed to Import.
- The **Deployment Status** modal now correctly displays the status of the deployment process.
- An error message is added that notifies you if validation errors are found for properties that are currently filtered.
- If available, the display names of configuration property keys are displayed above the property key.

Highly available Kafka Connect integration

SMM uses the Kafka Connect service role's REST URL to establish a connection with Connect and serve Connect metrics. Previously, even if your Connect deployment was highly available and had multiple service roles deployed, SMM could only be configured with a single connection URL. From now on, multiple URLs can be configured. If the Connect service role that SMM is connected to fails, SMM automatically connects to a different instance that is available.

As a result of this change, the Kafka Connect Host and Kafka Connect Port properties are replaced by the Kafka Connect Rest HostPorts property. If Kafka Connect Rest HostPorts is left empty (default), SMM is automatically

configured with the host, port, and protocol of the Connect service role instances belonging to the Kafka service selected with the Kafka Service SMM property.

If you previously configured Kafka Connect Host and Kafka Connect Port, the values set in the properties are automatically migrated to Kafka Connect Rest HostPorts when you upgrade.

What's New in Streams Replication Manager

Learn about the new features of Streams Replication Manager in Cloudera Runtime 7.2.17.

Improved SRM logging

SRM's logging capabilities are improved. From now on:

- Kafka clients created by SRM's internal connectors reference the replication flow they are a part of ([KAFKA-14838](#) backport).
- SRM now includes references to the replication flow in the log context of its internal connectors.

These changes enable differentiation between the logs associated with each replication flow.

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

Learn about the new features of Hadoop YARN and YARN Queue Manager in Cloudera Runtime 7.2.17.

Apache Hadoop YARN

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

YARN Queue Manager

Dynamic Queue Scheduling

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

For more information, see [Dynamic Queue Scheduling](#).

Queue priority

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

Setting Maximum Parallel Application Limits

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

For more information, see [Setting Maximum Parallel Application](#).

Editing placement rules

Support to edit previously created placement rules was added.

For more information, see [Editing placement rules](#).

Refresh queues option in Queue Manager UI

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

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

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

For more information, see [Configuring the resource capacity of root queue in absolute mode](#).

New YARN Queue Manager Overview Page

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

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

What's New in Apache ZooKeeper

Learn about the new features of ZooKeeper in Cloudera Runtime 7.2.17.

Added support to read password from file

Read key/trust store password from file (ZOOKEEPER-4396).

Unaffected Components in this release

There are no new features for the following components in Cloudera Runtime 7.2.17.

- Data Analytics Studio
- Apache Hadoop HDFS
- Apache Hive
- Apache Impala
- Apache Oozie
- Apache Solr

Fixed Issues In Cloudera Runtime 7.2.17

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.17.

Fixed Issues in Atlas

Review the list of Apache Atlas issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-54852: Backward compatibility for check provided for AttributeName in Parent and Child TypeDef

This patch provides backward compatibility for 2 changes mentioned: <https://issues.apache.org/jira/browse/ATLAS-3872> Restrict typedef creation when a child type attribute conflicts with parent type attribute of same name <https://issues.apache.org/jira/browse/ATLAS-4522> Updating typedef with new supertype should be allowed only if attributes are unique compared to other existing supertypes

CDPD-53717: Failed to extract ADLS Meta Data in Azure cdp_atlas_nat run

Fixed failed to extract ADLS Meta Data in Azure cdp_atlas_nat run

CDPD-50914: Atlas - Upgrade reactor-netty to 1.0.24+ due to CVE-2022-31684

Upgrade reactor-netty to 1.0.24+ due to CVE-2022-31684

CDPD-49792: Dynamic Index Recovery issues and improvements

Merged in CDH-7.2.17.x

CDPD-48720: The attributes of __AtlasAuditEntry type where not getting indexed in the Solr, because the entities of __AtlasAuditEntry gets created before the attributes gets indexed (TypeDef gets created).

This issue is resolved.

CDPD-49053: Atlas - Upgrade Tinkerpop to 3.5.4

Upgrade Tinkerpop to 3.5.4

CDPD-48166: Atlas - Upgrade snakeyaml due to CVE-2022-1471

Upgrade snakeyaml due to CVE-2022-1471

CDPD-48090: Atlas - Upgrade icu4j to 66.1+ due to CVE-2020-21913

Upgrade icu4j to 66.1+ due to CVE-2020-21913

CDPD-48034: Atlas - Upgrade jettison to 1.5.2 due to CVE-2022-45685 and CVE-2022-45693

Upgrade jettison to 1.5.2 due to CVE-2022-45685 and CVE-2022-45693

CDPD-48026: Atlas - Upgrade Woodstox to 5.4.0/6.4.0 due to multiple CVEs

Upgrade Woodstox to 5.4.0/6.4.0 due to multiple CVEs

CDPD-47988: Atlas - Upgrade Netty to 4.1.86.Final due to CVE-2022-41881, CVE-2022-41915

Upgrade Netty to 4.1.86.Final due to CVE-2022-41881, CVE-2022-41915

CDPD-47912: Atlas - Upgrade moment.js to 2.29.4 due to CVE-2022-24785, CVE-2022-31129

Updated moment.js version to 2.29.4

CDPD-46519: aws-s3-extractor.sh fails with ClassNotFoundException: com.fasterxml.jackson.core.util.JacksonFeature

Fixed aws-s3-extractor.sh Atlas script fails in 7.2.16 AWS test runs

CDPD-46409: Atlas - Upgrade Spring-security to 5.6.9/ 5.7.5 due to CVE-2022-31692, CVE-2022-31690

Upgrade Spring-security to 5.7.5 due to CVE-2022-31692, CVE-2022-31690

CDPD-45097: Atlas - Prevent re-indexing of deleted relationship edges to improve entity-update execution process

Implementing this fix along with fix in CDPD-36992 together would ensure that deleted relationship edges are not reattempted for deletion and re-indexing is not done for such edges. As a result, execution time for messages piling on the Atlas hook (Kafka lag) would be reduced drastically leading to significant performance improvement.

CDPD-44811: [PbC] Atlas - Replace log4j 1.x with log4j2

Merge to CDH 7.2.17 (7.2.17.x), pre-cdpd-master (7.2.17) branch,

CDPD-44104: Every hive insert generates an Atlas audit event

As per the current architecture, Atlas maintains audit data for all DML events also. By introducing "DML audit filter" feature, Atlas can skip all the DML event processing which improves performance by reducing unnecessary processing efforts from Atlas. Eventually, there will not be a need to store any such audit events which will reduce use of storage. The DML audit filter will be enabled by default with the below explained configuration. If the user wants to have audit data for DML events, it can be configured using the same configuration. Filtering DML events will be handled by Atlas hook, DML audit filter configuration should be enabled at Hive hook. Note that DML Audit filter is enabled by default, set to false to disable it Hive configuration Property :

Hive Service Advanced Configuration Snippet (Safety Valve) for atlas-application.properties
Configuration value : atlas.hook.hive.skip.dml.messages=true

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

Upgrade Spring Framework to 5.3.20 due to CVE-2022-22971, CVE-2022-22970

CDPD-36991: Backward compatibility for check provided for AttributeName in Parent and Child TypeDef

This patch provides backward compatibility for 2 changes mentioned: <https://issues.apache.org/jira/browse/ATLAS-3872> Restrict typedef creation when a child type attribute conflicts with parent type attribute of same name <https://issues.apache.org/jira/browse/ATLAS-4522> Updating typedef with new supertype should be allowed only if attributes are unique compared to other existing supertypes

CDPD-35438: When classification is created with multiple super types having same attributes , Atlas doesn't throw an exception

This patch restricts classification typedef creation when a child type attribute conflicts with parent type attribute of same name.

CDPD-31728: Audits : For db create , there are 2 update audits instead of 1 create 1 update

This patch will check for the Shell entities created and then assign the Entity status to the Audit entry.

CDPD-29307: Kafka producer entity stays in incomplete state in Atlas

The Kafka-Atlas plugin now fully creates Producer and Consumer entities and won't generate incomplete ones.

Apache patch information

- ATLAS-4746
- ATLAS-4576
- ATLAS-4743
- ATLAS-4736
- ATLAS-4733
- ATLAS-4727
- ATLAS-4719
- ATLAS-4622
- ATLAS-4735
- ATLAS-4679
- ATLAS-4655
- ATLAS-4668
- ATLAS-4335
- ATLAS-4666
- ATLAS-4464
- CDPD-56588 : [Atlas] test_purge_entity_api failed
- CDPD-56198: Regression: Atlas logs are not generated
- CDPD-55533: [Regression] hive_process and hive_process_execution (lineage) being generated for simple DML UPDATE queries run via hive
- CDPD-53364: Reinstate Atlas Docker images for 7.2.17
- CDPD-55282: Atlas - 7.2.17 - Update NOTICE files
- CDPD-53717 : Failed to extract ADLS Meta Data in Azure cdp_atlas_nat run
- CDPD-49792: Dynamic Index Recovery Improvements
- CDPD-44811: Atlas - Replace log4j 1.x with log4j2
- CDPD-50412: ATLAS-4733 : Download Basic and DSL search results
- CDPD-49302 : [PbC] Expose Iceberg tables in Atlas
- CDPD-49495: Fixed OOM Issue with DSL search caching

- CDPD-44104 : Every hive insert generates an Atlas audit event
- CDPD-54630 : ATLAS-4743: UI:Remove unused libraries from build package
- CDPD-53904: UI-Sync dashboardV2/Header.js with Apache master
- CDPD-50682: Fixed the Regression where Data Migration root URI was not getting redirected to the migration status page
- CDPD-48936: Fixed default value for atlas.entity.audit.differential to true
- CDPD-51876: ATLAS-4736 : Audit tab reporting a date of 1970 in a few properties.
- CDPD-50639 : Messages are sent to ATLAS_HOOK_UNSORTED but not seen in ATLAS_HOOK
- CDPD-48788 : Provide HA support for REST notification server
- CDPD-48789 : Provide Image support for REST notification server This commit does not contain secrets
- CDPD-48790 : PAM authentication not supported by REST notification server. This commit does not contain secrets
- CDPD-46359: HiveProcess - Output entity creation is ignored in case of same entity exists both inputs and outputs
- CDPD-48026 : Upgrade Woodstox to 5.4.0
- CDPD-47912 : (UI)Upgrade moment.js to 2.29.4 due to CVE-2022-24785, CVE-2022-31129 in 7.2.16.x
- CDPD-49053 : ATLAS-4728 - Upgrade Tinkerpop to 3.5.4
- CDPD-49393 : Atlas is failing in P2 phase for CDH 7.2.16.1
- CDPD-47599: import-hbase.sh script fails with UriBuilder Exception
- CDPD-47908 : UI:Upgrade Swagger to the latest version(Upgrade DOMPurify to the latest version to avoid Security issues): This commit does not contain secrets.
- CDPD-36991 : ATLAS-4576 Backward compatibility for check provided for AttributeName in Parent and Child TypeDef
- CDPD-48720: ATLAS-4727: admin/audits , admin/purge APIs fail with '[_AtlasAuditEntry.startTime] is not indexed in the targeted index [vertex_index]'
- CDPD-48684: Lineage is missing for CTAS tables created via impala-shell/beeline
- CDPD-47603: Added Atlas Preprocessor at SerialEntityProcessor to handle S3 V2 directory objectPrefix
- CDPD-46688 : Provide HA support for REST notification server
- CDPD-46687 : Provide Image support for REST notification server This commit does not contain secrets
- CDPD-48166 : ATLAS-4723 - Upgrade snakeyaml to 1.33
- CDPD-48034 : Atlas - Upgrade jettison to 1.5.2
- CDPD-48122 : admin/audits , admin/purge fail with '[_AtlasAuditEntry.startTime] is not indexed in the targeted index [vertex_index]'
- CDPD-47988 : Upgrade Netty to 4.1.86.Final
- CDPD-47173: Log4j migration: Replace DailyRollingFileAppender with RollingFileAppender
- CDPD-47813 : Atlas - Do the fix for CVE-2022-34271
- CDPD-46339 : ATLAS-4695:UI: Clicking on term assigned to deleted entity in entity details page results in error.
- CDPD-40822: Atlas - Upgrade Spring Framework in REST Notification module
- CDPD-47435 : ATLAS-4715 :(UI) Option to delete assigned term from deleted entity is visible
- CDPD-25084: Implemented concurrent ingest
- CDPD-46686 : PAM authentication not supported by REST notification server. This commit does not contain secrets
- CDPD-47307 : Term assigned to historical entity is not visible in Search results page
- CDPD-47356: REST Notification - messages from same source with same creation time are getting ignored
- CDPD-42014 : Upgrade jackson-databind to 2.12.7.1
- CDPD-46292 : Updating NOTICE file for cdpd-master
- CDPD-46519 : aws-s3-extractor.sh fails with ClassNotFoundException: com.fasterxml.jackson.core.util.JacksonFeature
- CDPD-46569 : Atlas - Upgrade Apache Ivy to 2.5.1 due to CVE-2022-37865, CVE-2022-37866
- CDPD-46492 : ATLAS-4710 : (UI)Relationship Attribute Filter button keeps loading
- CDPD-45331 : Embedded Server to host HTTP REST endpoint to handle notification messages This commit does not contain secrets

- CDPD-46409 : Upgrade Spring-security to 5.7.5
- CDPD-46089 : ATLAS-4677 - Atlas client lib throws - NoClassDefFoundError: org/apache/commons/configuration2/Configuration
- CDPD-43109 : ATLAS-4693,ATLAS-4661,ATLAS-4613 : (UI)Switching within Terms page does not land you in the same tab
- CDPD-43820 : Upgrade azure-storage-blob to 12.18.0
- CDPD-31728 : ATLAS-4666 : Intermittently, the audits for creation of hive_db registered are different than expected
- CDPD-46074: ATLAS-4689:UI: Basic Search: Invalid attributes passed in request
- CDPD-34249: REST Notification HOOK side - various hook topics patch
- CDPD-45763 : Atlas - Upgrade Apache Commons Text to 1.10.0 due to CVE-2022-42889
- CDPD-45671 :ATLAS-4691: Discrepancy in the atlas debug metrics between the active and the not active servers
- CDPD-45375 : Upgrade snakeyaml to 1.32
- CDPD-44695 : Regression : adls-extraction fails with ClassNotFoundException: com.fasterxml.jackson.core.util.JacksonFeature
- CDPD-35438 : ATLAS-4668 : When classification is created with multiple super types having same attributes , Atlas doesn't throw an exception
- CDPD-44508: ATLAS-4674: Regression: Classification tagging is not sending appropriate notification to ATLAS_ENTITIES
- CDPD-45576 : Upgrade jettison to 1.5.1
- CDPD-45221: ATLAS-4681: Relationship Search : SortBy is not working
- CDPD-45097 : Atlas kafka topic lag is very high and is not decreasing
- CDPD-35212: ATLAS:4558: Fix AtlasRepairIndex tool when tls is enabled

Fixed Issues in Avro

Review the list of Avro issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-47852: Get rid of old CDH versions and parent

Got rid of old CDH repository reference for artefacts to supports build with newer dependency versions

CDPD-45628: Avro - Upgrade Apache Maven to 3.8.6 due to CVE-2021-26291

Remove maven prerequisites of version 2.2.1, upgrade maven-core to 3.8.6 to fix CVE-2021-26291, and plexus-utils version to 3.5.0, apache file-management version to 3.0.0 to support the upgrade

Apache patch information

None

Fixed Issues in Cloud Connectors

Review the list of Cloud Connectors issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-46175: HADOOP-18521. ABFS prefetching input stream corruption

CDPD-56046: S3 distcp --update overwrites date of target file, even when names and content match

Changed test to match the product. Distcp updates the date of the file to the current timestamp if a different file is moved there.

CDPD-50522: Data-con setup fails, causing some tests to run with invalid configs

Fixed in CDPQE-16900 / <https://github.com/infracore/cloudera-qa/data-connectors-qa/commit/1d47c10b0f3eaffa0f95c5713dba31e848fca76b>

CDPD-48449: distcp -update skips files of same size, name when transferring from Hdfs to S3

The Distcp -update option, may encounter potential inaccuracies by skipping the copy when doing incremental update of files with identical names and sizes during the transfer process from HDFS to S3 or ABFS. This occurs due to the absence of checksum verification between the files for different stores. In order to address this concern, we employ the modification time as a means to minimize the occurrence of incorrect skips. If the source file has been modified more recently than its corresponding destination file, we proceed with the copy operation; otherwise, the file is skipped.

CDPD-46543: HADOOP-18526. Leak of S3AInstrumentation instances via hadoop Metrics references

HADOOP-18526. Leak of S3AInstrumentation instances via hadoop Metrics references

CDPD-45959: Some tests fail with ssl3_get_server_certificate:certificate verify failed

"fs.azure.ssl.channel.mode" has been set to "Default_JSSE". Switch to "Default" if the version of OpenSSL installed in your OS can successfully negotiate SSL connections with azure to achieve possibly improved performance.

Apache Patch Information

- HADOOP-18526
- HADOOP-18596
- HADOOP-18521

Fixed issues in Cruise Control

Review the list of Cruise Control issues that are resolved in Cloudera Runtime 7.2.17.

OPSAPS-66403: Cruise Control auth users lists are getting generated wrongly

Cruise Control handles the different authentication levels for the users correctly. When users are added to higher authentication levels, the lower level permissions are also assigned. For example, an ADMIN level user automatically has USER and VIEWER permissions as well.

Fixed issues in Data Analytics Studio

Review the list of Data Analytics Studio (DAS) issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-46997: Multiple unavailable views and test timeout regressions

Product change due to EOL disabled a functionality. Once reenabled, the affected tests ran fine.

CDPD-17000: DAS - Upgrade to JQuery 3.5.1

Solved as part of CDPD-35793, old jQuery removed.

Fixed Issues in Apache Hadoop

Review the list of Hadoop issues that are resolved in Cloudera Runtime 7.2.17.

Apache Patch Information

HADOOP-18602

Fixed Issues in HBase

Review the list of HBase issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-50683: HBase: Add idempotency support for fs.rename while using ABFS

HBaseAzureSemantics now supports resilient rename for files between source and destination on Azure Blob storage. However, HBase does not support any directory renaming and the behavior is same as ABFS.

Resilient rename is a workaround for `RenamePathFile 500/OAuthServerTimeoutError` and `RenamePathFile 404/OAuthClientOtherError` (Source File does not exist after a successfully rename) handled by the file system `HBaseAzureSemantics`. With this workaround, the region server must not crash during the commit stage of the flush operation for `HFile`.

CDPD-48344: Make the initial corePoolSize configurable for ChoreService

Add `hbase.choreservice.initial.pool.size` configuration property to set the initial number of threads for the `ChoreService`.

CDPD-48185: Hbase_mcc - Upgrade Scala to 2.12.5+/2.13.9+ due to CVE-2017-15288

The Scala version depends on the spark-scala version.

CDPD-46164: HBASE Tarball does not contain JWT related JARS

JWT related JARS added to the HBASE Tarball.

CDPD-46065: Job to sync schema differences between the HA clusters

A new tool `SyncTableDescriptorsTool` which synchronizes the table descriptors between two peers that uses the HBase multiple cluster client after one of the primary or the failover cluster disconnected from the two-way peer replication, especially, if the operator finds any table attributes that are different on the clusters.

To use this tool:

1. Get the classpath including `hbase-mcc.jar`, HBase's jars, and `hbase-site.xml` from the source cluster.
2. Use the following command:

```
java -cp "hbase-mcc-0.2.0-SNAPSHOT.jar:/path/to/clusters/://opt/cloudera/parcels/CDH/jars/*:`hbase classpath`" com.cloudera.hbase.mcc.SyncTableDescriptorsTool -t <table that needs to sync for table descriptor>
```

Apache Patch Information

- HBASE-27565

HBase new features:

- HBASE-27347: Port `FileWatcher` from ZK to autodetect keystore/truststore changes in TLS connections
- HBASE-27346: Autodetect key/truststore file type from file extension
- HBASE-27280 Add mutual authentication support to TLS (#4796)
- HBASE-27565 Make the initial `corePoolSize` configurable for `ChoreService` (#4958)
- HBASE-27506 Optionally disable sorting directories by size in `CleanerChore` (#4896)

HBase improvements:

- HBASE-27713 Remove `numRegions` in `Region Metrics` (#5107)
- HBASE-27726 Handling of ruby shell `SyntaxError` exceptions (#5147)
- HBASE-26526 Introduce a timeout to shutdown of WAL (#3297)
- HBASE-27684: add client metrics related to user region lock. (#5081) (#5133)
- HBASE-27646 Should not use `pread` when prefetching in `HFilePreadReader` (#5063) (#5123)
- HBASE-27710 `ByteBuff` ref counting is too expensive for on-heap buffers (#5115)
- HBASE-27670 Improve `FSUtils` to directly obtain `FSDataOutputStream` (#5064)
- HBASE-27458 Use `ReadWriteLock` for region scanner readpoint map (#5068)
- HBASE-15242: add client side metrics for timeout and remote exceptions. (#5023) (#5054)
- HBASE-21521 Expose master startup status via web UI (#4788) (#5021)
- HBASE-27626 Suppress noisy logging in `client.ConnectionImplementation` (#5019)
- HBASE-27551 Add config options to delay assignment to retain last region location (#4945)
- HBASE-27556 Reuse Zookeeper session of Master in `LogCleaner` (#4946)

- HBASE-27474 Evict blocks on split/merge; Avoid caching reference/hlinks if compaction is enabled (#4868)

HBase bugfixes:

- HBASE-27820: HBase is not starting due to Jersey library conflicts wi... (#5210) (#5261)
- HBASE-27752: Update the list of prefetched files upon region movement (#5194) (#5222)
- HBASE-27810. Check if event processor is already shut down (#5212)
- HBASE-27768 Race conditions in BlockingRpcConnection (#5154)
- HBASE-27445 fix the result of DirectMemoryUtils#getDirectMemorySize (#4846)
- HBASE-27778 Incorrect ReplicationSourceWALReader.totalBufferUsed may cause replication hang up (#5163)
- HBASE-27704 Quotas can drastically overflow configured limit (#5153)
- HBASE-27758 Inconsistent synchronization in MetricsUserSourceImpl (#5149)
- HBASE-27750: Update the list of prefetched Hfiles upon block eviction (#5140)
- HBASE-26866 Shutdown WAL may abort region server (#4254)
- HBASE-27676 Scan handlers in the RPC executor should match at least one scan queues (#5074)
- HBASE-27736 HFileSystem.getLocalFs is not used (#5125)
- HBASE-27671 Client should not be able to restore/clone a snapshot after it has TTL expired it's TTL has expired (#5118)
- HBASE-27651 hbase-daemon.sh foreground_start should propagate SIGHUP and SIGTERM
- HBASE-27718 The regionStateNode only need remove once in regionOffline (#5106)
- HBASE-27714 WALEntryStreamTestBase creates a new HBTU in startCluster method which causes all sub classes are testing default configurations (#5101)
- HBASE-27686: Recovery of BucketCache and Prefetched data after RS Crash (#5080)
- HBASE-27688 HFile splitting occurs during bulkload, the CREATE_TIME_TS of hfileinfo is 0 (#5097)
- HBASE-27673 Fix mTLS client hostname verification (#5065)
- HBASE-24781 Clean up peer metrics when disabling peer (#4997)
- HBASE-27650 Merging empty regions corrupts meta cache (branch-2) (#5038)
- HBASE-27668 PB's parseDelimitedFrom can successfully return when there are not enough bytes (#5059)
- HBASE-27644 Should not return false when WALKey has no following KVs while reading WAL file (#5032)
- HBASE-27661 Set size of systable queue in UT (#5053)
- HBASE-27636 The "CREATE_TIME_TS" value of the hfile generated by the HFileOutputFormat2 class is 0 (#5034)
- HBASE-27648 CopyOnWriteArrayMap does not honor contract of ConcurrentMap.putIfAbsent (#5031)
- HBASE-27643 [JDK17] Add-opens java.util.concurrent (#5028)
- HBASE-27619 Bulkload fails when trying to bulkload files with invalid names after HBASE-26707 (#5014)
- HBASE-27590 Change Iterable to List in SnapshotFileCache (#4995)
- HBASE-27621 Also clear the Dictionary when resetting when reading compressed WAL file (#5016)
- HBASE-27602 Remove the impact of operating env on testHFileCleaning (#5003)
- HBASE-27580 Reverse scan over rows with tags throw exceptions when using DataBlockEncoding (#5006)
- HBASE-26967 FilterList with FuzzyRowFilter and SingleColumnValueFilter evaluated with operator MUST_PASS_ONE doesn't work as expected(#4820)
- HBASE-27547 Close store file readers after region warmup (#4942)
- HBASE-25516 [JDK17] reflective access Field.class.getDeclaredField("modifiers") not supported (#3443)
- HBASE-27579 CatalogJanitor can cause data loss due to errors during cleanMergeRegion (#4986)
- HBASE-27561 hbase.master.port is ignored in processing of hbase.masters (#4952)
- HBASE-27529 Provide RS coproc ability to attach WAL extended attributes to mutations at replication sink (#4924)
- HBASE-27560 fix consistencyCheck did not report the hole on last region (#4950)
- HBASE-27540 add client side counter metrics for failed rpc calls (#4929)
- HBASE-27524 Fix python requirements problem (#4918)
- HBASE-27519 Another case for FNFE on StoreFileScanner after a flush followed by a compaction (#4923)

- HBASE-27498: Added logic in ConnectionImplementation.getKeepAliveMasterService to avoid expensive rpc calls in synchronized block (#4889)
- HBASE-27491 Do not clear cache on RejectedExecutionException (#4914)
- HBASE-27487 Addendum remove unused imports
- HBASE-27487: Slow meta can create pathological feedback loop with multigetts (#4900)
- HBASE-27494: Fix missing meta cache dropping exception metrics (#4902)
- HBASE-27484 FNFE on StoreFileScanner after a flush followed by a compaction (#4882)
- HBASE-27503 Support replace <FILE-PATH> in GC_OPTS for ZGC (#4892)
- HBASE-26320 Implement a separate thread pool for the LogCleaner (#4895)
- HBASE-27463 Reset sizeOfLogQueue when refresh replication source (#4863)
- HBASE-27504 Remove duplicated config 'hbase.normalizer.merge.min_region_age.days' in hbase-default.xml (#4894)
- HBASE-27464 In memory compaction 'COMPACT' may cause data corruption when adding cells large than maxAlloc(default 256k) size (#4881)
- HBASE-27043 Let lock wait timeout to improve performance of SnapshotHFileCleaner (#4437)
- HBASE-25899 Improve efficiency of SnapshotHFileCleaner (#3280)
- HBASE-25363 Improve performance of HFileLinkCleaner by using ReadWriteLock instead of synchronize
- HBASE-27379 fix numOpenConnections metric is one less than the actual (#4884)
- HBASE-27001 The deleted variable cannot be printed out (#4883)
- HBASE-27469 IllegalArgumentException is thrown by SnapshotScannerHDFSACLController when dropping a table (#4865)
- HBASE-27414 Adjust hfilelink alternative paths order (#4847)
- HBASE-27450 Update all our python scripts to use python3 (#4851)
- HBASE-27440 fix table HistogramMetrics leak in table metrics map (#4838)
- HBASE-27159 Emit source metrics for BlockCacheExpressHitPercent (#4830)
- HBASE-27339 Improve sasl connection failure log message to include server (#4823)
- HBASE-27401 Addendum remove unused imports
- HBASE-27406 Make /prometheus endpoint accessible from HBase UI (#4833)

HBase connectors bugfixes:

- HBASE-27630: hbase-spark bulkload stage directory limited to hdfs only (#108)
- HBASE-27624 Cannot Specify Namespace via the hbase.table Option in Spark Connector (#107)
- HBASE-27397 Spark-hbase support for 'startWith' predicate (#105)

Technical Service Bulletins

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

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

Fixed Issues in HDFS

Review the list of HDFS issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-33801: Backport HDFS-16420 which fixes a data corruption bug running balancer on EC data.

Fixed bug where in rare cases Erasure Coded blocks may be removed permanently.

Apache Patch Information

HDFS-16420

Fixed Issues in Apache Hive

Review the list of Hive issues that are resolved in Cloudera Runtime 7.2.17.

OPSAPS-67031: Increase the hive split threads to 64 for AWS and GCP

This fix addresses a performance issue for Ranger Raz with specific cloud providers. As part of this fix, the default value for Hive split processing threads (`hive.compute.splits.num.threads`) is increased to 64. This change affects only Amazon Web Services (AWS) and Google Cloud Platform (GCP) deployments.

CDPD-39232: Performance impact caused by RANGER-3593 when running SHOW TABLES query

During authorization checks, the Ranger plugin looks for ownership information for database objects (database, tables, or views) through HMS API calls if the ownership information is not provided by Hive. For databases having a large amount of tables, this can result in slow performance for statements like SHOW TABLES because the Ranger plugin has to generate an API call for each object.

This fix ensures that ownership information is provided by Hive for the above statements so that the Ranger authorization plugin does not have to issue HMS API calls.

CDPD-40779: Hive - Upgrade netty to 4.1.77 due to CVE-2022-24823

Upgraded netty to 4.1.77 to fix CVEs.

CDPD-43490: Hive Security - Upgrade jackson-databind to 2.12.7.1 due to critical CVEs

Upgraded jackson-databind to 2.12.7.1 to fix CVEs.

CDPD-43509: Hive Security - Upgrade dom4j: flexible XML framework for Java to safe version due to critical CVEs

Removed dom4j to fix CVEs.

CDPD-46097: Null Pointer Exception with mask udf

This fix addresses the issue where the 'SHA512' masked value is not being propagated to Tez executors.

CDPD-46360: Authorize DataConnectors in Hive until CDPD-26882 is implemented

Data connectors in HIVE are temporarily authorized using RWSTORAGE privileges until CDPD-26882 is resolved.

CDPD-46568: Hive - Upgrade Apache Ivy to 2.5.1 due to CVE-2022-37865, CVE-2022-37866

Upgraded apache ivy to 2.5.1 to fix CVEs.

CDPD-46664: Hive - Upgrade commons-codec to 1.13 or higher

Upgraded commons-codec to 1.15 to fix CVEs.

CDPD-47132: Pushdown Date data type to metastore via direct SQL/ JDO

Fix partition filtering when querying partition metadata from Hive Metastore and the partition key column data type is 'date'.

CDPD-47464: ALTER VIEW command is allowed even when user has a deny policy on the underlying table

The ALTER VIEW statement was not authorized correctly. This fix addresses the security concern related to the authorization of ALTER VIEW AS queries.

CDPD-48022: Hive - Upgrade postgresql to 42.5.1 due to CVE-2022-41946

Upgraded PostgreSQL to 42.5.1 to fix CVEs

CDPD-48801: Pushdown Timestamp data type to metastore via direct SQL / JDO

Support partition filtering when querying partition metadata from Hive Metastore and partition key column data type is 'timestamp'.

CDPD-49145: Oozie and Spark tests are failing in multi-comp-pre with ZooKeeper-based or direct JDBC URL to Hive

HiveServer (HS2) uses the `InetAddress.getHostName()` API to get its host name and register itself with ZooKeeper. The API behaviour is changed on JDK 11 with specific operating systems and returns only the host name without the domain suffix. Therefore, HiveServer is not accessible to clients when the server information is obtained from Zookeeper.

To address this issue, the `InetAddress.getCanonicalHostName()` API is used to get the host name along with the fully qualified domain name.

CDPD-49507: {OWNER} policy not working with HIVE UDFs in RangerHiveAuthorizer

The UDFs used in Hive will now honor {Owner} policies in ranger with this fix.

CDPD-50450: Inconsistency between session Hive and thread-local Hive may cause HS2 deadlock

Two HS2 sessions can go into a deadlock state and can indefinitely wait for each other (related to RANGER-3593). This fix resolves the deadlock condition.

CDPD-55168: HiveConnection: HTTP Response code: 404 Failed to connect to master node on YCloud

This issue has been fixed based on the support provided in Cloudera Manager for Knox Custom Topology management.

CDPD-55914: SELECT query on table with remote database returns NULL values with postgresQL and Redshift data connectors

This fix addresses the issue where some datatypes are not mapped from postgresQL or Redshift to Hive data types in the connector, which results in displaying null values for the columns of these data types.

CDPD-56133: Compaction entry dequeue order

This fix addresses the issue where compaction requests are dequeued in a random order. With this fix, the compaction requests are now dequeued in a First In First Out (FIFO) order.

Apache Patch Information

- HIVE-25313
- HIVE-26594
- HIVE-26625
- HIVE-26681
- HIVE-26753
- HIVE-26778
- HIVE-26787
- HIVE-26850
- HIVE-26914
- HIVE-27116
- HIVE-27147
- HIVE-27201
- HIVE-27285
- HIVE-27316
- HIVE-27330

Fixed Issues in Hive Warehouse Connector

There are no fixed issues for HWC in Cloudera Runtime 7.2.17.

Fixed Issues in Hue

Review the list of Hue issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-48893: Hue logs get overwritten

In previous implementations, multiple file handlers would write to a single log file, causing the Hue logs to be overwritten. Hue now uses a socket handler, which solves this problem.

CDPD-48787: Hue does not load after setting the download_bytes_limit

Earlier, when you set the value of the download_bytes_limit property in the [beeswax] section in the Hue Advanced Configuration Snippet, the Hue web interface did not load with the following error: Processing exception: unsupported operand type(s) for /: 'str' and 'int'. This issue has been fixed.

CDPD-48281: Add limit request field configs to hue

New configs have been added to control the request header limits used by Gunicorn server.

CDPD-48250: Import table data to external location fails when % is included in the filename

Earlier, when you tried to upload a file containing the “%” symbol using the Hue File Browser, the “%” symbol was encoded incorrectly. You would see an unexpected symbol in the filename. This issue has been resolved by fixing the encoding issue in the Hue importer.

CDPD-47077: Issues while exporting or importing documents in Hue

Earlier, you would have faced issues while exporting and importing workflows and documents in Hue. This was due to a bug in Django 1.11, which has been fixed in all Django versions higher than 2.2. In the CDP 7.2.17 release, Hue uses Django 3.2, which has automatically resolved this issue.

CDPD-46957: Export feature is not available from search bar in Hue

This issue has been fixed. You now see an option to download document results in the top search bar and left assist pane in Hue.

Fixed Issues in Apache Impala

Review the list of Impala issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-54464: IMPALA-12043 Large catalog info triggers "TTransportException: MaxMessageSize reached"

Fix "TTransportException: MaxMessageSize reached" during large catalog info reading.

CDPD-50862: Extend usage of fire_listener_event API to HS2/Spark to generate events on DML queries

The data written from Spark will now generate an "Insert" event in the HMS notification log table. This is useful for external services (listeners) like Impala to know the current condition of an external table.

CDPD-50186: IMPALA-11966 Enable cache_ozone_file_handles by default

Enables cache_ozone_file_handles by default to improve scan performance with Ozone.

CDPD-50180: IMPALA-11920 Spill to HDFS/Ozone can't address by service name

Specifying an HDFS or Ozone path for spilling to external storage can now use a service name, such as ofs://ozone1/myvolume/spilldir/.

CDPD-49648: Upgrade chart.js to 2.9.4+ due to CVE-2020-7746

Upgrades chart.js in Impala UI to address CVE-2020-7746.

CDPD-49100: Impala - Upgrade Kryo to 5.0.2+ due to BDSA-2016-1151(DoS, memory corruption, and RCE attacks)

Kryo is not used by Impala at runtime. It is excluded from the Impala Java build to avoid CVE concern.

CDPD-49015: IMPALA-11859 Metric tracking encrypted bytes read

Adds BytesReadEncrypted to query profiles and the metric `impala-server.io-mgr.encrypted-bytes-read` to observe reads of encrypted data from HDFS/Ozone.

CDPD-48780: impala-shell now requires setuptools be manually added

Fixes a regression in CHF3 where `impala-shell` under Python 2 required installing `setuptools`.

CDPD-48721: Impala - Upgrade JQuery Datatables to the latest version to avoid Security issues

Updates JQuery Datatables in the Impala UI to address CVE-2020-28458 and CVE-2021-23445.

CDPD-47643: Impala SHOW statement to display EC files and policies

Impala's `SHOW FILES`, `SHOW PARTITIONS`, `SHOW TABLE STATS`, and `DESCRIBE EXTENDED` now display the erasure code policy for files/tables in filesystems that support erasure coding.

CDPD-47640: Impala erasure coding support

Impala now supports interacting with erasure-coded files in HDFS.

CDPD-47206: IMPALA-11730 Add support for spilling to Ozone

Impala can now be configured to spill to Ozone, for example with `scratch_dirs="/tmp/scratch,ofs://ozone-scm:9862/tmp"`.

CDPD-47205: IMPALA-11736 LOAD DATA statement with Ozone data can not load data from different bucket

Impala now correctly handles loading data from a different Ozone bucket.

CDPD-47030: Impala-shell ldap_password_cmd fails on Python 3.8

Fixes `impala-shell --ldap_password_cmd` with Python 3.

CDPD-46986: Backport CDPD-45163 to CDH-7.2.16.1 branch

Implement Iceberg manifest caching config for Impala

CDPD-46985: Backport Iceberg manifest caching to CDH-7.2.16.1

Add Iceberg manifest caching feature for Impala.

CDPD-46368: Impala remote Ozone scans slow even after data cache warmup

Improves scan performance with Ozone when using a data cache.

CDPD-45886: IMPALA-11670 Upgrade components for CVEs, make it easier to override versions

Upgrades components `guava` and `jackson-databind` to address CVE-2020-8908, CVE-2022-42003, and CVE-2022-42004.

CDPD-45661: Support erasure-coding in impala

Reading erasure-coded files from Ozone is now supported with Impala.

CDPD-44372: Impala - Upgrade Spring Framework to 5.3.20 due to multiple CVEs

Impala upgrade the Spring framework to 5.3.20 to address multiple CVEs: - CVE-2022-22971 - CVE-2022-22968 - CVE-2022-22970

CDPD-43746: Support for Ozone erasure coded data

Impala now supports interacting with erasure-coded files in Ozone.

CDPD-24718: Application is Vulnerable to CSRF attack - Impala CatalogServer, StateStore

Adds Cross-Site Request Forgery (CSRF) prevention in Impala UIs. Changing log levels now requires a POST request from the UI or with the 'X-Requested-By' custom header.

CDPD-8130: Add HTTP Strict Transport Security (HSTS) for Impala

Adds HTTP Strict Transport Security (HSTS) to Impala UI responses when HTTPS is enabled.

Apache Patch Information

- IMPALA-11892
- IMPALA-9487

- IMPALA-11755
- IMPALA-12031
- IMPALA-11476
- IMPALA-7003
- IMPALA-11913
- IMPALA-12037
- IMPALA-11856
- IMPALA-11704
- IMPALA-8518

Fixed Issues in Apache Kafka

Review the list of Apache Kafka issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-29307: Kafka producer entity stays in incomplete state in Atlas

The Kafka-Atlas plugin now fully creates producer and consumer entities and does not generate incomplete ones.

CDPD-48822: AvroConverter ignores default values when converting from Avro to Connect schema

The AvroConverter now propagates field default values to Connect schemas.

OPSAPS-65485: Selecting the Require Connectors To Override Kafka Client JAAS Configuration property causes automatic Kafka Connect startup retries to fail

Kafka Connect does not fail on start retries when the Require Connectors To Override Kafka Client JAAS Configuration property is selected.

CDPD-53179: Amazon S3 Sink connector fails when buffer size is reached

The Amazon S3 Sink connector no longer fails when there is more than 5 MB (buffer size) of data available in a Kafka source topic and the connector receives more than 5 MB of data in a single poll.

Apache patch information

- KAFKA-14838: Add flow/connector/task/role information to MM2 Kafka client.id configs

Fixed Issues in Apache Knox

Review the list of Knox issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-55168: HiveConnection: HTTP Response code: 404 Failed to connect to master node on YCloud

Once both CM and CDH changes re: OPSAPS-66676 were available, the issue was solved

CDPD-53722: Knox - Upgrade OkHttp to 3.14.9/4.10.0 due to medium CVEs - PvC

Upgrade OkHttp to 3.14.9/4.10.0 due to medium CVEs.

CDPD-51895: CM discovery should consider only the configured discovery username and password , should not fallback to default admin user credentials

CM discovery will not use default admin user credentials when discovery credentials are explicitly configured but are not valid.

CDPD-49983: Atlas/Ranger/CM/webhdfs api via Knox Proxy fails with "java.io.IOException: Close SendCallback@a12ceb4[PROCESSING]"

Fixed by virtue of CB-20899

CDPD-49206: Refine should perform discovery check

Service discovery is not triggered is service URL is missing.

CDPD-48847: Oozie "root" rewrite rule's pattern is too open

<https://issues.apache.org/jira/browse/KNOX-2841>

CDPD-48241: Knox - Upgrade mina to 2.1.5+ due to CVE-2021-41973

Upgrade mina to 2.1.6 due to CVE-2021-41973

CDPD-48021: Knox - Upgrade postgresql to 42.5.1 due to CVE-2022-41946

Upgraded postgresql to 42.5.1 to address CVE-2022-41946

CDPD-47749: Knox - Upgrade Spring Framework to 6.0.0 due to CVE-2016-100027

Knox is not affected by this vulnerability.

CDPD-47037: RM UI redirect link to the Spark3 History Server fails

Spark 3 History Server link Resource Manager UI works in Yarn UI v1 too.

CDPD-46666: Knox - Upgrade commons-codec to 1.13 or higher

Upgraded commons-codec due to CVE

CDPD-46560: Knox - Upgrade protobuf-java to 3.16.3/3.19.6/3.20.3/3.21.7 due to CVE-2022-3171

Upgrade protobuf-java to 3.16.3/3.19.6/3.20.3/3.21.7 due to CVE-2022-3171.

CDPD-45349: Server Side Request Forgery - Knox - Host Parameter

KNOX-2827 Dispatch whitelist regular expression is matched against the base URL

CDPD-42463: Knox - Upgrade OkHttp to 3.14.9/4.10.0 due to medium CVEs

Upgrade OkHttp to 3.14.9/4.10.0 due to medium CVEs.

CDPD-42153: Knox - Upgrade Protocol Buffer Java API to 2.6.1/3.21.2 due to medium CVEs

Upgrade Protocol Buffer Java API to 2.6.1/3.21.2 due to medium CVEs.

CDPD-41897: Knox - Upgrade Bouncy Castle to 1.70 due to medium CVEs

Upgrade Bouncy Castle to 1.70 due to medium CVEs.

Apache patch information

- KNOX-2871
- KNOX-2827
- KNOX-2841
- KNOX-2911

Fixed Issues in Apache Kudu

Review the list of Apache Kudu issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-47036: Upgrade elasticsearch (a transitive dependency through Ranger) to 7.17.1+ due to CVEs

Upgraded elasticsearch to 7.17.1+ due to CVEs.

CDPD-47068: Update default value for --tablet_history_max_age_sec to avoid OOM for kudu-master

Fixed an issue with the kudu-master process consuming too much memory in case of very large clusters, clusters with many thousands of tables, or clusters with huge numbers of DDL operations per day.

The default setting of --tablet_history_max_age_sec for tablet servers has been set to 7 days since logical backup/restore has been implemented with CDH6.3.0 release (corresponding to upstream Kudu release 1.10.0), but storing that much history for system catalog doesn't make much sense as Kudu masters don't scan the system tablet with timestamps back in the past. For bigger Kudu clusters with many nodes and thousands of tables, if there is a sustainable high rate of DDL activity or Raft election storms in the cluster, the system tablet might accumulate a lot of deltas of its history. That might lead to kudu-master processes using a lot of memory while performing rowset merge compactions.

CDPD-53423: Kudu was not able to connect to Ranger in public cloud due to an incorrect environment variable name

Kudu control script was using the wrong environment variable to fill Ranger service name. This caused the ranger subprocess to crash. As a result, all the Kudu CLI commands failed due to absence of ranger subprocess on the other end to cater CLI requests. As a fix, the correct environment variable is used to ensure a valid ranger service name is picked when Kudu cluster is set up.

KUDU-3450: Buffer was too low for messages between Kudu and the Ranger client, causing authorization failures when the access control lists were too large

The issue is now fixed.

Kudu HMS Sync is disabled and is not yet supported

The issue is now fixed.

Kerberos authentication fails with rdns disabled

When rdns is set to false, the Kudu Java client does not retain the original hostname, and replaces them with the resolved IP addresses. This prevents Kerberos authentication from working properly. For more information, see [KUDU-3415](#).

The issue is now fixed.

Apache Patch Information

KUDU-3406

Fixed Issues in Apache Livy

Review the list of Apache Livy issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-49189: Ordering and pagination support in livy GET /statement request

Fixed the code by adding the missing API documentation and the code logic to display session statements as latest-first based on the query parameter.

CDPD-48733: Merge Apache Livy 0.7.2 into CDP DataHub 7.2.17

Livy and Livy for Spark 3 have been updated to upstream version 0.7.2. LDAP is not supported.

CDPD-48118: Provide ttl field for a livy session

Added the support to provide TTL (time to leave) field for Livy session API.

CDPD-47766: Return session information with livy sessions APIs

Enhanced the Livy session APIs to return new fields.

Apache Patch Information

- LIVY-968
- LIVY-967
- LIVY-970

Fixed Issues in Apache Oozie

Review the list of Oozie issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-26975: Using the ABFS / S3A connectors in an Oozie workflow where the operations are "secured" may trigger an IllegalArgumentException with the error message java.net.URISyntaxException: Relative path in absolute URI.

This issue is fixed now.

CDPD-48847: Oozie "root" rewrite rule's pattern is too open

The "root" rewrite rule for Oozie is too open and causes the following issue:

When you use Apache Hue as the UI for Oozie and define a workflow property for a file path like `hdfs://mnameservice1/oozie/test`, then on Hue's workflow details page, you will see an url for that property like: `http://oozie-host.examole.com:11000/oozie/test`

The issue is now fixed.

Apache patch information

None

Fixed Issues in Ozone

There are no fixed issues for Ozone in Cloudera Runtime 7.2.17.

Apache patch information

Apache patches in this release.

- None

Fixed Issues in Phoenix

Review the list of Phoenix issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-26858: Phoenix timezone handling to be fixed

Phoenix now supports the `phoenix.query.applyTimeZoneDisplacement` property, which enables a more consistent and standards-compliant handling of `java.sql.Date/Time/TimeStamp` types. The property is set as `FALSE` by default for backwards compatibility, and it must be set to `TRUE` explicitly for the new behaviour. Using the `java.time.LocalDate/LocalTime` types are also directly supported according to the JDBC 4.2 specifications.

See [PHOENIX-5066](#) and [PHOENIX-6881](#) for more details.

Apache Patch Information

- PHOENIX-5066

Calcite Avatica new features:

- [CALCITE-4536] Add support for BIT data type (Zeng Rui)

Calcite Avatica improvements:

- CALCITE-5581 Implement Basic client side load balancing in Avatica Driver
- CALCITE-5327 Make SSL key-store type configurable
- CALCITE-5218 Verify HTTP client class before instantiating it
- CALCITE-4877 Make the exception information more explicit for instantiate plugin.
- CALCITE-4757 Allow columns of type Null in ResultSet (NobiGo)
- CALCITE-3401 Assume empty keystore passwords by default (Istvan Toth, Alessandro Solimando)

Calcite Avatica bugfixes:

- CALCITE-4752 PreparedStatement#SetObject() fails for BigDecimal values
- CALCITE-4971 Update httpClient and httpcore to latest 5.1 release
- CALCITE-5009 Transparent JDBC connection re-creation may lead to data loss
- CALCITE-4962 Protobuf debug does not show request/response type
- CALCITE-4152 Upgrade Avatica to use the configurable SPNEGO Jetty implementation
- CALCITE-4828 Standard exception console output

- CALCITE-4837 FLOOR/CEIL for DECADE, CENTURY, MILLENIUM return wrong results
- CALCITE-4573 NullPointerException while fetching from a column of type ARRAY
- CALCITE-4602 ClassCastException retrieving from ARRAY that has mixed INTEGER and DECIMAL elements
- CALCITE-4600 ClassCastException retrieving from an ARRAY that has DATE, TIME or TIMESTAMP elements
- CALCITE-4767 Add Quoting.BACK_TICK_BACKSLASH (Jack Scott)
- CALCITE-4752 PreparedStatement#SetObject() fails for BigDecimal values
- CALCITE-4676 Avatica client leaks TCP connections
- CALCITE-4503 Order of fields in records should follow that of the SQL types (Alessandro Solimando)
- CALCITE-3163 Incorrect mapping of JDBC float/real array types to Java types (Ralph Gasser)
- CALCITE-3881 DateTimeUtils.addMonths yields incorrect results (Zhenghua Gao)
- CALCITE-4476 DateTimeUtils.timeStringToUnixDate may produce wrong time (Vladimir Ozerov)
- CALCITE-4181 Avatica throws exception when select field is a List<Object> (Kent Nguyen)
- CALCITE-4379 Meta.Frame created with java float values in rows hits a ClassCastException in toProto()
- CALCITE-4196 Consume all data from client before replying with HTTP/401
- CALCITE-4138 Metadata operations via Avatica turn empty string args to null

Phoenix improvements:

- PHOENIX-6944 Randomize mapper task ordering for Index MR tools
- PHOENIX-6873 Use cached Connection in IndexHalfStoreFileReaderGenerator
- PHOENIX-6395 Reusing Connection instance object instead of creating everytime in PhoenixAccessController class (addendum: don't close the shared Connection on stop)
- PHOENIX-6881 Implement the applicable Date/Time features from JDBC 4.2
- PHOENIX-6834 Use Pooled HConnection for Server Side Upsert Select

Phoenix bugfixes:

- PHOENIX-6969 Projection bug in hinted uncovered index query with order by
- PHOENIX-6953 Creating indexes on a table with old indexing leads to inconsistent co-processors
- PHOENIX-6874 Support older HBase versions with broken ShortCircuitConnection
- PHOENIX-6872 Use ServerUtil.getConnection() in UngroupedAggregateRegionScanner
- PHOENIX-6395 Reusing Connection instance object instead of creating everytime in PhoenixAccessController class
- PHOENIX-5066 The TimeZone is incorrectly used during writing or reading data
- PHOENIX-6823 calling Joda-based round() function on temporal PK field causes division by zero error
- PHOENIX-6889 Improve extraction of ENCODED_QUALIFIERS
- PHOENIX-6720 CREATE TABLE can't recreate column encoded tables that had columns dropped
- PHOENIX-6855 Upgrade from 4.7 to 5+ fails if any of the local indexes exist.
- PHOENIX-6818 Remove dependency on the i18n-util library to fix CVE
- PHOENIX-6841 Depend on omid-codahale-metrics
- PHOENIX-5894 Table versus Table Full Outer join on Salted tables not ... (#1395)

Phoenix connectors bugfixes:

- PHOENIX-6667: Spark3 connector requires that all columns are specified when writing

Phoenix Omid improvements:

- OMID-239: OMID TLS support (#129)

Phoenix Query Server improvements:

- PHOENIX-6810 Make SSL key-store type configurable

Phoenix Query Server bugfixes:

- PHOENIX-6908 KerberosName\$NoMatchingRule exception in QueryServer.PhoenixRemoteUserExtractor

- PHOENIX-6704 sqlline-thin.py doesn't work with python3
- PHOENIX-6762 Phoenix QueryServer cannot run correctly with python 3.8+

Fixed Issues in Parquet

Review the list of Parquet issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-47864: Parquet - CVE-2021-41561-Parquet is vulnerable to Dos attack.

Handle negative values in page headers that fixes CVE-2021-41561

Apache Patch Information

Fixed Issues in Apache Ranger

Review the list of Ranger issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-56554: [7.2.17 CLONE] - Turning usersync debug logging on results in users not getting synced due to NPE

Fix NPE while logging debug messages

CDPD-56213: Fix sql patch 65 syntax issue for oracle db

Fix sql patch 65 syntax issue for oracle db

CDPD-55997: Log4j2 support : Write java patches logs to log file

Log4j2 support : Write java patches logs to log file

CDPD-55994: Ranger Upgrade to 7.1.9 may fail

Fix for ranger upgrade failure

CDPD-55281: Ranger - 7.2.17 - Update NOTICE files

Notice file updated

CDPD-55164: ranger policy replication transform step is not printing logs

Improve ranger policy replication transformation logs

CDPD-54624: [CLONE] - Ranger - Upgrade commons-codec to 1.15

Upgrade commons-codec to 1.15

CDPD-53826: Ranger - Upgrade jettison to 1.5.4 due to CVE-2023-1436

Upgrade jettison to 1.5.4

CDPD-53805: Ozone_key tag based policies are not working

What was the Root Cause? Ozone qualified name parsing had a issue wherein '/' was getting included in the key name which resulted in wrong key matching while enforcing policy How was this Issue Resolved? Logic for parsing ozone qualified name changed such that '/' is not included in the key name which was causing issue previously.

CDPD-53804: Ranger - Upgrade Spring Framework to 5.3.26/6.0.7 due to CVE-2023-20861 and CVE-2023-20860

Upgrade Spring Framework to 5.3.27

CDPD-53440: Ranger audit metrics deletion is failing

Bug fixed for Ranger audit metrics deletion.

CDPD-50720: Regression caused by CDPD-45891

What was the issue? getDeletedGroups() was using incorrect URI. How was the issue fixed? Fix uri for getDeletedGroups() in PolicyMgrUserGroupBuilder

CDPD-50450: Backport HIVE-27201: Inconsistency between session Hive and thread-local Hive may cause HS2 deadlock

Two HS2 sessions can go into a deadlock state with RANGER-3593 and can indefinitely wait for each other. This patch resolves the deadlock condition.

CDPD-50395: Ranger - Upgrade org.json to 20230227+ due to CVE-2022-45688

Removed org.json dependency from Ranger KMS. Ranger KMS does not required this as direct dependency. org.json will be fetch as run time dependency for service Ranger KMS KTS.

CDPD-50299: Ranger - Upgrade Kerby to 2.0.3 due to CVE-2023-25613

Upgrade Kerby to 2.0.3

CDPD-50025: [7.2.17.0] Ranger: upgrade tomcat to 8.5.85 or higher

Upgrade tomcat to 8.5.86

CDPD-50019: Update servicedef by name results in 400 status code while the same request works with update servicedef using id

Update servicedef by name API should return http response code 200

CDPD-49882: [7.2.17 CLONE] - Ranger AD User Sync - support for AD group names containing slashes

Support for LDAP/AD usernames and group names with special chars

CDPD-49774: Tags (classification) are not getting synced when we Add attribute values for classification

Sync Tags (classification) when attribute values for classification is added

CDPD-49711: assignPermissionToUser in XUserMgr creates entries with NULL moduleId in x_user_module_perm

Fixed assignPermissionToUser in XUserMgr to correct the bug which assigns permissions for a module (which does not exist) to users with Auditor role.

CDPD-49701: Ranger S3 policy fails for non recursive access to the root of a bucket

There was an issue in Ranger S3 policies for 7.2.16 Raz-enabled environments where policies with Allow conditions and only read or write were not being honored (i.e., users will be denied) by Ranger and policies with only Deny conditions and only read or write were not being honored (i.e., users will be allowed) by Ranger. This jira fixed this issue and the Ranger S3 policies in Raz-enabled environments would be honored.

CDPD-49651: Ranger Tagsync - Convert to Web Application

This task converts the Ranger Tagsync module into a Web Application.

CDPD-49589: Add yarn and impala users to audit filter for solr servicedef to avoid logging of audits

Add yarn and impala users to audit filter for solr servicedef

CDPD-49588: clientIP is not logged for create/grant/revoke role operations via hive beeline

What was the root cause? Server.getRemoteIp() does not return client ip correctly. How was the issue fixed? Use SessionState to log clientIP in RangerHiveAuthorizer.

CDPD-49373: Groups are not visible in mask and row level policy listing tables.

Fixed that all groups are listed properly in policy listing table

CDPD-48947: Ranger Upgrade from 7.2.11 to 7.2.16 failed

Fix for Ranger upgrade failure from source version CDH-7.2.11

CDPD-48828: [ranger] [replication] Script should not permit N : 1 mappings for services of the same service type

Restrict duplicate mappings to same source service name in ranger replication configuration

CDPD-48394: Ranger is opening a lot of zk connections when solr is down

Making sure that Ranger closes the Zookeeper connection in case when Solr service is not reachable. Also following the configured number of retries to connect to Solr and on given time intervals.

CDPD-48389: Change in api response for get APIs

Ranger REST API response object will not include properties/fields which are NULL or empty/
blank. RANGER-3948

CDPD-48337: exportJson api returns all policies in repo when filter string used has reponame and groupName

exportJson api should return ranger policies as per filter specification

CDPD-48322: queryparams repositoryType and groupName are not working for /service/public/api/policy api

Fix for failing get policy API with queryparams

CDPD-48232: [ranger] [replication] Policy transform step is removing hdfs execute permission.

Keep hdfs execute permission during Policy transformation

CDPD-48165: Ranger - Upgrade snakeyaml due to CVE-2022-1471

Upgrade snakeyaml to 2.0

CDPD-48129: [ranger][replication] If a change is made in the resource field of a policy on the source cluster, a new policy is created on the target cluster instead of changing the existing policy

Add support to get matching ranger policies by given algorithm

CDPD-48119: Ranger - Upgrade OWASP Java HTML Sanitizer due to security CVEs

Upgrade OWASP Java HTML Sanitizer

CDPD-48041: Ranger - Upgrade commons-net to 3.9.0 due to CVE-2021-37533

Upgrade commons-net to 3.9.0

CDPD-48032: Ranger - Upgrade jettison to 1.5.2 due to CVE-2022-45685 and CVE-2022-45693

Upgrade jettison to 1.5.2

CDPD-47994: [Ranger] Not able to fetch Policy details using guid /api/policy/guid/{guid} without service name

Fix for failing get policy by GUID API

CDPD-47989: Ranger - Upgrade Netty to 4.1.86.Final due to CVE-2022-41881, CVE-2022-41915

Upgrade Netty to 4.1.86.Final

CDPD-47909: Ranger - Upgrade moment.js to 2.29.4 due to CVE-2022-24785, CVE-2022-31129

Upgrade moment.js to 2.29.4 due to CVE-2022-24785, CVE-2022-31129

CDPD-47900: Log4j2 support in Ranger

Log4j 1.x dependency is removed and upgraded to log4j2

CDPD-47856: Ranger - Upgrade bootstrap to 6.0.0 due to GHSA-87mg-h5r3-hw88

Upgrade bootstrap to 5.5.3

CDPD-47760: [Ranger][UserSync]Enumerate Groups will give error on executing 'getent group' command

What was the issue? incorrect usage of getent command in UnixUserGroupBuilder How was the issue fixed? Fixed the usage of getent in UnixUserGroupBuilder

CDPD-47464: Alter view command allowed even when user has a deny policy on the underlying table

"Alter View As" queries were not being authorized correctly. This patch addresses the security concern around the authorization of "Alter View As" queries.

CDPD-47056: Fix Ranger TagRest API deleteTagResourceMapByGuid

Fix Ranger TagRest API deleteTagResourceMapByGuid

CDPD-46961: [aws][7.2.7->7.2.16] solr-server error after the DL upgrade

Modified default_value column type to TEXT of x_service_config_def table.

CDPD-46866: [cdpd-master clone] - Ranger - Upgrade Woodstox to 5.4.0/6.4.0 due to multiple CVEs

Upgrade Woodstox to 5.4.0

CDPD-46789: Policy update request fails if isDenyAllElse flag is set true in request json when using "/policy/apply" API

Fix for Policy update request failure when isDenyAllElse flag is set to true in in "/policy/apply" API request json

CDPD-46781: Restrict scripts from accessing Java classes and methods

Improve validation of condition expressions used in Ranger policies.

CDPD-46677: Ranger - Upgrade Woodstox to 5.4.0/6.4.0 due to multiple CVEs

Upgrade Woodstox to 5.4.0

CDPD-46667: Ranger - Upgrade commons-codec to 1.13 or higher

Upgrade commons-codec to 1.14

CDPD-46659: Ranger - Upgrade wildfly-openssl to 1.1.3.Final/1.1.3.Final+ due to CVE-2020-25644

Upgrade wildfly-openssl to 1.1.3.Final

CDPD-46561: Ranger - Upgrade protobuf-java to 3.16.3/3.19.6/3.20.3/3.21.7 due to CVE-2022-3171

Upgrade protobuf-java to 3.21.7 to fix a CVE issue

CDPD-46447: [CR-7.2.17] Add 'preload' directive to HSTS header

Successfully added preload directive in HSTS i.e. Strict-Transport-Security tag in response header.

CDPD-46256: Ranger Audit metrics page broken in New UI

Fixed Audit metrics not loading in New UI

CDPD-46244: [CR-7.2.16] Add 'preload' directive to HSTS header

Successfully added preload directive in HSTS i.e. Strict-Transport-Security tag in response header.

CDPD-46233: Knox plugin is not working

Knox service was failing when Audit metrics was enabled. Fix was done to handle the CNF error in Knox Ranger plugin which took care of this error

CDPD-46161: [ranger][replication] cm_hdfs service wasn't transformed properly

Fix for hdfs service policies transformation failure

CDPD-46160: [ranger][replication] Export should fail for non-existing services

Ranger policy export should fail for non existing services

CDPD-46097: NPE during ranger ctas masking test

This patch addresses the issue where the "SHA512" masked value is not being propagated to Tez executors.

CDPD-45533: AuditFileSpool logs out all events that were not audited successfully

AuditFileSpool should log only those events which are audited successfully

CDPD-44997: Upgrade snakeyaml to 1.32 in ranger-plugins-audit

Upgrade snakeyaml library to 1.32 to fix a CVE issue

CDPD-44645: Investigate alternative for enumerate=true in SSSD conf

Document usersync configs for using FreeIPA

CDPD-44513: HA support for Ranger TagSync

HA support for Ranger TagSync added as part of this new feature enhancement.

CDPD-43640: HA support for Ranger User Sync

HA support for Ranger UserSync added as part of this new feature enhancement.

CDPD-43132: Allow roles, tagrest & xaudit Ranger Admin APIs via Knox proxy

This fix allows access to ranger role, tagrest and xaudit ranger admin APIs from Knox proxy.

CDPD-43037: Add/ Update metric details for Ranger TagSync

This new feature provides Application specific metrics and JVM metrics details for Ranger Tagsync module.

CDPD-41446: Create common Ranger HA module

Common HA module created as part of this new feature enhancement.

CDPD-40964: Need to update Knox re-write rules to allow access to newer APIs introduced in Ranger

Allow metrics,roles, tagrest & xaudit Ranger Admin APIs via Knox proxy

CDPD-39608: RANGER : [cdpd-master] Upgrade Jackson-core and Jackson-databind due to CVE[2020-36518]

Successfully upgraded Jackson-core to v2.12.7. and databind to v2.12.7.1

CDPD-38189: Make sure that Ranger plug-in can insert audit documents when Solr is upgraded in rolling fashion

When Solr is in Rolling upgrade, plugin audits will be stored in local filesystem when Solr is not able to reach at any point of time and will be pushed when available.

CDPD-30591: Provide option to update group memberships when same users/groups are synced from different sync sources

Allow sync source updates for existing users synced via different sync sources

CDPD-29102: Ranger - Remove log4j 1.x dependencies due to EOL

Log4j 1.x dependency is removed and upgraded to log4j2

CDPD-15744: HA support for Ranger Tag Sync/User Sync

HA support for Ranger TagSync and UserSync added as part of this new feature enhancement.

Apache Patch Information

- RANGER-4241
- RANGER-4242
- RANGER-3821
- RANGER-4163
- RANGER-4173
- RANGER-4159
- RANGER-4135
- RANGER-4204
- RANGER-4113
- RANGER-4112
- RANGER-4115
- RANGER-4153
- RANGER-4131
- RANGER-4073
- RANGER-4109
- RANGER-3947
- RANGER-4205
- RANGER-4031
- RANGER-3975
- RANGER-3991
- RANGER-4028
- RANGER-4043
- RANGER-3977
- RANGER-4206

- RANGER-3995
- RANGER-3959
- RANGER-3962
- RANGER-3957
- RANGER-3961
- RANGER-4151
- RANGER-4150
- RANGER-4149
- RANGER-3729
- RANGER-3498
- RANGER-4148

Fixed Issues in Schema Registry

Review the list of Schema Registry issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-48568: JAR storage does not work on AWS S3 for Schema Registry

Schema Registry Amazon S3 JAR storage now functions correctly.

CDPD-49217 and CDPD-50309: Schema Registry caches user group membership indefinitely

Schema Registry now evicts Kerberos user and group information from its cache with a configurable time.

CDPD-54379: KafkaJsonSerializer and KafkaJsonDeserializer do not allow null values

The KafkaJsonSerializer and KafkaJsonDeserializer now properly translates null payloads as null.

CDPD-48822: AvroConverter ignores default values when converting from Avro to Connect schema

The AvroConverter now propagates field default values to Connect schemas.

CDPD-48888: Schema Registry generates redundant schemas when byte[] with default field exists

Schema Registry's schema normalization and fingerprinting mechanism has been enhanced to properly handle default values for bytes data types.

CDPD-53380: Schema Registry Client should retry the request on Knox gateway errors

The Schema Registry Client will retry Knox gateway related failed requests as defined by the request retry configuration.

CDPD-48853: Schemas created with the Confluent Schema Registry API cannot be viewed in the UI

Schemas created in Cloudera Schema Registry using the Confluent Schema Registry API are now visible in the Cloudera Schema Registry UI.

In addition, the `/api/v1/schemaregistry/search/schemas/aggregated` endpoint of the Cloudera Schema Registry API now correctly returns schemas created with the Confluent Schema Registry API.

Fixed Issues in Apache Solr

Review the list of Solr issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-50032: Solr: CVE-2023-24998-upgrade commons-fileupload library to version 1.5

Backport upstream jira SOLR-14250 and SOLR-14461 which removes using commons-fileupload (and uses jetty instead).

CDPD-44607 and CDPD-46198: Upgrade jsoup to 1.15.3 to fix CVE-2021-37714 and CVE-2022-36033

Upgraded jsoup version as part of CVE fix.

CDPD-45967: Upgrade hsqldb to 2.7.1 due to CVE-2022-41853

Upgraded HSQLDB version as part of CVE fix.

Apache Patch Information

None

Fixed Issues in Spark

Review the list of Spark issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-55243: Fix case sensitivity of Iceberg's CachingCatalog

Previously, using inconsistent casing for database and table names of Iceberg tables in queries can lead to Spark reading a stale cached snapshot after a write to the table (append, update, delete) in the same Spark session. Now the cache is insensitive to the case of database and table names and is always refreshed on a write in the session.

CDPD-50862: Extend usage of fire_listener_event API to HS2/Spark to generate events on DML queries

The data written from Spark will now generate an "Insert" event in the HMS notification log table. This is useful for external services (listeners) like Impala to know the current condition of an external table.

CDPD-46530: Spark - Upgrade Apache Ivy to 2.5.1 due to CVE-2022-37865, CVE-2022-37866

Apache Ivy upgraded to 2.5.1 to avoid CVE

CDPD-46306: CVE-2022-31777: Apache Spark XSS vulnerability in log viewer UI Javascript

[SPARK-39505][UI] Escape log content rendered in UI

Apache patch information

None

Fixed Issues in Apache Sqoop

Review the list of Sqoop issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-44431: Disable the Sqoop direct mode feature with ability to enable it again temporarily

Sqoop's direct mode is no longer supported and is disabled by default. However, you can still enable it by either setting the `sqoop.enable.deprecated.direct` property globally in Cloudera Manager for Sqoop or by specifying it in the command-line through `-Dsqoop.enable.deprecated.direct=true`.

CDPD-44531: Sqoop cannot export Parquet data due to ClassCastException

Sqoop can now export the following data types from Avro and Parquet files:

- Int, Float, Double to the same RDBMS types
- Long to BigDecimal, Date, Time, TimeStamp
- Bytes to BigDecimal
- Fixed to Decimal and TimeStamp

Note that Fixed to TimeStamp does not work if the source date is based on the Julian calendar.

CDPD-50423: Sqoop ClassCastExceptions when exporting from Parquet

This fix introduces enhancements in Sqoop when exporting from Parquet. Additional data type mappings are now supported.

CDPD-52721: Replace log4j 1.x with reload4j

log4j has been replaced with reload4j in Sqoop.

Apache patch information

None

Fixed Issues in Streams Messaging Manager

Review the list of Streams Messaging Manager issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-46728: SMM UI shows the consumerGroup instead of the instances on the Profile page's right hand side

The **Consumer Group Profile** page now correctly shows the consumer instances on the right hand side. Previously the consumer groups were shown.

CDPD-46465: Searching for workers on the connector overview page freezes the page

Using the search field on the **Connect Cluster Profile** tab no longer freezes the page.

CDPD-45406: The Connector Profile page of unassigned connectors is blank

The **Connector Profile** page of unassigned connectors are now correctly rendered and display that the connector is in an unassigned status.

CDPD-46073: Data Explorer loads indefinitely

The **Data Explorer** page no longer breaks if the partition parameter is manually removed from the URL.


CDPD-26633: The SMM API returns SMTP passwords of email notifiers in its response

The /notifiers endpoint of the SMM API no longer returns the SMTP password in its responses.

CDPD-49227: The Cluster Replications page crashes if the co-located cluster unknown to SRM

The **Cluster Replications** page is now correctly displayed even when the co-located Kafka cluster is unknown to SRM.

CDPD-56086: The Data Explorer modal displays the messages of the wrong topic

The **Data Explorer** modal that you open by clicking  on the **Topics** page now displays the messages of the selected topic.

CDPD-49696: Certain alerts may crash the Alerts page

Composite alerts with one of the conditions containing an assertion on cluster metrics no longer crashes the UI.

Fixed Issues in Streams Replication Manager

There are no fixed issues for Streams Replication Manager in Cloudera Runtime 7.2.17.

Fixed Issues in Apache Tez

Review the list of Tez issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-49216: Tez - Upgrade jquery-ui to 1.13.0+ due to CVEs

jquery-ui 1.13.0 version we already delivered into cdpd-master and cdw-master branches in the part of <https://jira.cloudera.com/browse/CDPD-35798>. Closing this Jira and we will handle jquery-ui 1.13.2 version upgrade in the part of <https://jira.cloudera.com/browse/CDPD-44443>

CDPD-48031: Tez - Upgrade jettison to 1.5.3 due to CVE-2022-45685 and CVE-2022-45693

Upgraded the jettison version to 1.5.3 to fix CVEs

CDPD-40867: Tez - Upgrade gson to 2.9.0 due to CVE-2022-25647

Tez is not using direct dependency of gson. This dependency is injected from hadoop-common:jar:3.1.1.7.2.15.4 If we want to change in gson version, we need to change hadoop-common version(which is gson version 2.9.0+ used. hadoop fixed gson CVE issues as part of <https://jira.cloudera.com/browse/CDPD-40852>) Marking as resolved as gson version is dependent

on HADOOP. This dependency is not been used directly. If any additional information about this issue from tez side please let me know. Thanks !

Apache patch information

None

Fixed Issues in Apache YARN and YARN Queue Manager

Review the list of YARN and YARN Queue Manager issues that are resolved in Cloudera Runtime 7.2.17.

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

Fixed the issue so logs are now visible.

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

Fixed issue.

COMPX-14340: YARN-11490 JMX QueueMetrics breaks after mutable config validation in CS

Fix: JMX metrics broke after 2 or more configuration validation.

COMPX-14147: YARN-11312 [UI2] Refresh buttons don't work after EmberJS upgrade

Fixed the issue of EmberJs variable change from targetObject to _targetObject

COMPX-14120: Backport YARN-11463: Node Labels root directory creation doesn't have a retry logic

Retry logic is implemented and backported for root directory creation during RM node label store initialization.

COMPX-13773: YARN-11461 NPE in determineMissingParents when the queue is invalid

Fix NPE log warning when submitting to invalid queue.

COMPX-13423: MAPREDUCE-7433 Remove unused mapred/LoggingHttpResponseEncoder.java

Unused code.

COMPX-13272: HADOOP-18602 Remove netty3 dependency

netty3 is removed from the dependencies

COMPX-12661: YARN-11075 Explicitly declare serialVersionUID in LogMutation class

The serialVersionUID field is explicitly set for the LogMutation class.

COMPX-12645: Backport YARN-10946

Changes are submitted to cdpd-master.

COMPX-12560: Fix JSON formatting in the RM tests

The previous backports are extended in cdpd-master, CDH-7.1.7.2000, CDH-7.2.16.x and CDH-7.1.8.x branches.

COMPX-12524: Backport YARN-10005

Code improvements in MutableCSConfigurationProvider

COMPX-11487: RM service experiences failure at recovery when auto created queue has issues

Fix NPE log warning when submitting to invalid queue.

COMPX-11384: Backport YARN-11063

Fixed Issue Text

COMPX-10909: Investigate if placement rules are working fine if username contains dot, and default queue is set to that queue

Username with dot now will work well with CS placement rules

COMPX-7460: YARN-10965. Introduce universal capacity resource vector

New feature added to CS, not yet wired in so it won't change any existing CS config or behaviour.

CDPD-50485: [7.2.17][azure] yarn applicaitons stuck in accepted state

Root cause fixed in <https://jira.cloudera.com/browse/OPSAPS-66526>.

COMPX-13719: Queue creation is failing with 500 errors in cdp_cloud_7_2_17

Fixed as part of OPSAPS-66613

COMPX-13567: MAPREDUCE-7434 Fix ShuffleHandler tests

Unit test fix.

COMPX-13422: MAPREDUCE-7268 Fix TestMapreduceConfigFields

Test fix.

COMPX-13421: MAPREDUCE-7237 Supports config the shuffle's path cache related parameters

ShuffleHandler's cache related parameters are configurable.

COMPX-12964: MAPREDUCE-7431 ShuffleHandler is not working correctly in SSL mode after the Netty 4 upgrade

Bugfix.

COMPX-12922: QueueManager should allow configuring absolute capacities beyond current cluster capacity

n/a

COMPX-12872: [Config-Service]- Queue Manager - Upgrade OkHttp to 3.14.9/4.10.0 due to medium CVEs

Upgrade OkHttp to 4.10.0

COMPX-12803: QM 7.1.8 CHF4 - Upgrade Apache Commons Text to 1.10.0 due to CVE-2022-42889

n/a

COMPX-12487: Queue Manager 7.2.17 - Upgrade jackson-databind to 2.13.3 due to high CVEs

n/a

COMPX-12374: [CR-7.2.16.0] Notice file update - config-service, cpx

n/a

COMPX-12341: CPX [QM]Upgrade snakeyaml to 1.33 due to high CVEs

Upgrade snakeyaml to 1.33

COMPX-12340: CPX [Config-Store]Upgrade snakeyaml to 1.33 due to high CVEs

Upgrade snakeyaml to 1.33

COMPX-12246: Config Service - Upgrade Apache Commons Text to 1.10.0 due to CVE-2022-42889

Upgrade commons-text to 1.10.0

COMPX-12228: QM - Upgrade jersey to 2.35 / 3.0.2 due to CVE-2021-28168

Upgrade jersey to 2.37 and jackson to 2.13.4

COMPX-11479: [CPX] Queue Manager - Upgrade OkHttp to 3.14.9/4.10.0 due to medium CVEs

Removed unused OkHttp dependency

COMPX-11112: Queue Manager - Upgrade Bouncy Castle to 1.70 due to medium CVEs

n/a

Apache patch information

- YARN-11490
- YARN-11312
- YARN-11463
- YARN-11461

- YARN-11063
- YARN-11190
- MAPREDUCE-7433
- MAPREDUCE-7434
- MAPREDUCE-7268
- MAPREDUCE-7237
- MAPREDUCE-7431

Fixed Issues in Zeppelin

Review the list of Zeppelin issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-53819: Increase default Zeppelin RPC connection pool size based on ZEPPELIN-5005

Exposed `zeppelin.interpreter.connection.poolsize` and made it configurable as a safety value in the Zeppelin configuration file.

CDPD-48033: Zeppelin - Upgrade jettison to 1.5.2 due to CVE-2022-45685 and CVE-2022-45693

Added to 7.1.7 SP2 CHF4.

CDPD-41910: ZooKeeper - Upgrade Bouncy Castle to 1.70 due to medium CVEs

Upgraded Bouncy Castle to 1.70 due to critical CVEs.

Apache patch information

- None

Fixed Issues in Apache ZooKeeper

Review the list of ZooKeeper issues that are resolved in Cloudera Runtime 7.2.17.

CDPD-56215: Backport ZK client change to read password from file

Zookeeper is now able to Read Key/trust store password from file

CDPD-56134: Reload4j migration error in ZooKeeper

Fixed ZooKeeper log4j JMX issue with reload4j. <https://issues.apache.org/jira/browse/ZOOKEEPER-3737>

CDPD-30427: Fix custom ZooKeeper trust manager for FIPS

I checked the issue on a FIPS cluster in-house, but the error didn't show up. We can reopen the ticket later if needed.

Apache Patch Information

- ZOOKEEPER-4396
- ZOOKEEPER-3737
- ZOOKEEPER-4393

Fixed Issues In Cloudera Runtime 7.2.17.100

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.17.100.

CDH

- HOTREQ-1422 Need HIVE-26779 on Public Cloud runtime 7.2.15.10-1
- HOTREQ-1429 Request to backport CDPD-55677 into the next 7.2.15.x maintenance release

- HOTREQ-1506 Schema Registry schema import must not deduplicate schemas

CFM

- HOTREQ-1502 Ship CFM-3498 to 7.2.16 (CFM-2.2.6) and 7.2.17 (CFM-2.2.7)
- HOTREQ-1503 Ship CFM-3513 for 7.2.16 (CFM-2.2.6) and 7.2.17 (CFM-2.2.7)
- HOTREQ-1486 Ship CFM-3256 to 7.2.17 (CFM-2.2.7)
- HOTREQ-1459 Ship NIFI-11653 Security fix for 7.2.17 (CFM-2.2.7), 7.2.16 (CFM-2.2.6) and 7.2.15 (CFM-2.2.5)
- HOTREQ-1481 Ship NIFI-11744 security fix for 7.2.17 (CFM-2.2.7), 7.2.16 (CFM-2.2.6) and 7.2.15 (CFM-2.2.5)
- HOTREQ-1489 Ship NIFI-11560 to CFM-2.2.7 to fix one UI issue
- HOTREQ-1488 Ship NIFI-11334 to CFM-2.2.6.0 (7.2.16) and CFM-2.2.7.0 (7.2.17)

Known issue: CDPD-54714

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

Follow the procedure available in the [Integrate Hue with Knox](#) documentation.

Fixed Issues In Cloudera Runtime 7.2.17.200

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.17.200.

Cloudera Manager

- HOTREQ-1557 - Fix to Cloudera Manager health test failing - MGMT_PAUSE_DURATION

CDH

- HOTREQ-1478 Backport CDPD-46799 and CDPD-57996 to 7.2.16.xxx
- HOTREQ-1543 Backport CDPD-60778 to CDP 7.2.16 and 7.2.17
- HOTREQ-1525 Backport Python 3 fixes for `collect_minidumps.py` to 7.2.17 service pack
- HOTREQ-1559 HotFix for HIVE-27643
- HOTREQ-1515 Need HIVE-13288 on top of CDP 7.2.16

CFM

- HOTREQ-1540 Ship NIFI-11924 to 7.2.16 and 7.2.17
- HOTREQ-1539 Ship NIFI-11744 security fix for 7.2.16 and 7.2.17
- HOTREQ-1520 Ship NIFI-11854 for 7.2.16 (CFM-2.2.6) and 7.2.17 (CFM-2.2.7)
- HOTREQ-1541 Ship CFM-3471 to PCR-7.2.17
- HOTREQ-1531 Deliver NIFI-11617 to 7.2.17 (CFM-2.2.7.0)

CEM

- HOTREQ-1568 Ship CEM 1.6.0.0 to PCR-7.2.17

Fixed Issues In Cloudera Runtime 7.2.17.300

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.17.300.

CDH

- HOTREQ-1571 Add configuration option to make container allocation prefer nodes without reserved containers
- HOTREQ-1440 Query failures with error "repeated binary array (STRING) is not a group"
- HOTREQ-1578 Need Hotfix which includes CDPD-62091 and CDPD-62034
- HOTREQ-1590 Need hotfix for HIVE-21100 on CDP 7.2.15.0
- HOTREQ-1599 Fix for calculation of replication-records-lag in SRM
- HOTREQ-1612 Hotfix needed for HIVE-25576 in CDP-PublicCloud-DataHub-7.2.x clusters
- [TSB 2023-704](#): File corruption when downloading files larger than 1 MB from ABFS with Hue File Browser
- [TSB 2023-703](#): Risk of Data Loss when using Hue S3 File Browser

CFM

- HOTREQ-1587 SHIP NIFI-12160 for 7.2.17, 7.2.16 and 7.2.15
- HOTREQ-1597 Ship NIFI-11621 for 7.2.17 (CFM-2.2.7)

Fixed Issues In Cloudera Runtime 7.2.17.400

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.17.400.

CDH

- HOTREQ-1645 Request for Emergency HOTFIX to address TSB-746 in 7.2.17.x

CFM

- HOTREQ-1626 Ship NIFI-12596 to CFM-2.2.7.0 - 7.2.17
- HOTREQ-1655 Ship CFM CSD ranger fix to 7.2.17 and 7.2.18

Known Issues In Cloudera Runtime 7.2.17

You must be aware of the known issues and limitations, the areas of impact, and workaround in Cloudera Runtime 7.2.17.

Known Issues in Apache Atlas

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

CDPD-: 53176: Partition Specification data for Iceberg Table is not sent to Atlas in Hook context.

When a Iceberg table is created with partition spec, partition specification data is not sent to Atlas in Hook context. The partition specification data is stored differently for Hive than for Spark and Impala.

For example, for Spark and Impala, the partition data is present in Table parameters.default-partition-spec but for Hive partition data is stored in Partition Transform Information and not Table parameters.default-partition-spec. In case of Hive, Atlas is not getting Partition Transform Information or Table parameters.default-partition-spec from Hook context.

CDPD-: 56594: Lineage (spark_process) is not created for views created on Iceberg tables.

A non-Iceberg table creates lineage connecting the source table and view.

CDPD-: 59413: Plugin is not supported with older Atlas server versions for Iceberg tables.

Copy the model file 1130-iceberg_table_model.json to the directory: /opt/cloudera/parcels/CDH/lib/atlas/models/1000-Hadoop.

Proceed to restart the Atlas Service using Cloudera Manager.

CDPD-: 56590: Create table "like" from Iceberg table creates a hive_table instead of iceberg_table.

By default, for tables created using the "like" command, lineage is not generated in Atlas. The destination like table should be of the same type as source table. Instead an iceberg_table for source and hive_table for destination are getting created.

CDPD-: 56587: Lineage (spark_process) is not created for INSERT OVERWRITE / INSERT INTO SELECT Iceberg tables.

Querying on non-Iceberg tables using spark.sql creates a lineage (spark_process) connecting the source and the insert overwrite table. The same queries on an Iceberg tables using spark.sql does not create any lineage (spark_process).

CDPD-48122: Operations like admin/audits, admin/purge fail with a 500 internal server error message "[_AtlasAuditEntry.startTime] is not indexed in the targeted index [vertex_index]"

None

DOCS-17056: Atlas Rolling Upgrade related to Zero Downtime Upgrade (ZDU).

Upgrade process comprises of upgrading Cloudera Manager + Runtime upgrade + Operating System upgrade. Though Atlas cannot comply with a full ZDU process, there is no data loss observed through the entire upgrade process. Post upgrade, all the created entities before and during the upgrade process are available without any changes or modifications.

Some limitations that are observed during the ZDU process:

- While Atlas goes through the process of rolling upgrade, some downtime might be expected because Atlas does not support Active-Active model. Failover consumes sometime since Active-Passive is the currently supported model. As the Passive instance becomes Active, there is some downtime where Atlas is not reachable and the messages from clients are queued up in Kafka.
- Solr does not support Rolling Upgrade due to which Atlas REST requests fail during the Solr upgrade.
- Nodes unavailable due to OS Upgrade: Due to nodes going down and services not being accessible. (Not limited to Atlas but to also other available services).

CDPD-54964: ICEBERG External Table via impala-shell appears as hive_table, instead of iceberg_table.

Instead of appearing as iceberg_table, the entity appears as hive_table. Although the table parameters has ICEBERG data.

CDPD-55301: The ddlQueries and ALTERNATIVE_* lineage are missing for Spark tables created using spark3-shell.

The ddlQueries and outputFromProcesses (lineage) is missing for the alter queries.

CDPD-54990: The in-place migration of Hive table to Iceberg table with ALTER TABLE storage_handler using Beeline creates new iceberg_table entity but retains the old hive_table entity as is

Running the query results with Atlas having two entities with same name but different types. One with hive_table and another with iceberg_table.

CDPD-40346: The ddlQueries and ALTERNATIVE_ADDCOLS lineage missing for Impala tables.

The ALTERNATIVE_ADDCOLS lineage has some issue when an Impala table is altered and the corresponding lineage is not created.

CDPD-55671: When one Atlas server host is not reachable (stopped), the GET request does multiple failover for approximately 4 minutes and takes around 2 minutes for every failover and finally the request fails.

None

CDPD-55122: Any user with ssh access can view the downloaded results.

None

CDPD-57549: Rolling upgrade / ZDU: Atlas throws 503 when Zookeeper goes through upgrade

When Zookeeper goes through Rolling upgrade, Atlas REST calls throws 503 error. Entities created using Atlas Kafka hook are created in Atlas and no data loss is expected.

CDPD-46606: Performing Hive queries renders a notification for update data in the Hive table.

None

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

None

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

None

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

None

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

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

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

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

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

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

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

None

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

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

None

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

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

None

CDPD-11940: Database audit record misses table delete

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

None

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

In normal operation, Atlas receives metadata to create entities from multiple services on the same or separate Kafka topics. In some instances, such as for Spark jobs, metadata to create a table entity in Atlas is triggered from two separate messages: one for the Spark operation and a second for the

table metadata from HMS. If the process metadata arrives before the table metadata, Atlas creates a temporary entity for any tables that are not already in Atlas and reconciles the temporary entity with the HMS metadata when the table metadata arrives.

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

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

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

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

None

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

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

None

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

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

None

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

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

None

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

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

None

CDPD-9095: Duplicate audits for renaming Hive tables

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

None

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

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

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

None

CDPD-7781: TLS certificates not validated on Firefox

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

None

CDPD-6675: Irregular qualifiedName format for Azure storage

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

None

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

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

None

CDPD-4762: Spark metadata order may affect lineage

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

None

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

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

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

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

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

None

CDPD-3160: Hive lineage missing for INSERT OVERWRITE queries

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

None

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

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

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

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

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

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

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

The free text search bar in the top of the screen allows you to search across entity types and through all text attributes for all entities. The search shows the top 5 results that match the search terms at

any place in the text (*term* logic). It also shows suggestions that match the search terms that begin with the term (term* logic). However, in this release, the search results are case-sensitive.

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

CDPD-1823: Queries with ? wildcard return unexpected results

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

None

CDPD-1664: Guest users are redirected incorrectly

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

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

CDPD-922: IsUnique relationship attribute not honored

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

None

CDPD-24058: The Atlas-Kafka hook creates a new entity instead of linking them

When the import-kafka.sh tool is used and later the plugin is enabled in Kafka configurations, new incomplete topic entities is created. The tool is not linking the existing topics with the clients.

None

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

The error occurred when Schema Registry tried to make a relationship in Atlas between a schema and a non-existent corresponding topic.

Known Issues in Apache Avro

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

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

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

None.

Known Issues in Cloud Connectors

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

CDPD-48113: The option key to enable/disable readahead on the ABFS connector is changed to fs.azure.enable.readahead.v2 It is enabled by default, so read-ahead is active even if fs.azure.enable.readahead is false. This ensures that readahead is automatically re-enabled on releases where the HADOOP-18521 is fixed.

None

CDPD-46175: HADOOP-18521. ABFS prefetching input stream corruption

set fs.azure.enable.readahead to false

Known issues in Cruise Control

There are no known issues for Cruise Control in Cloudera Runtime 7.2.17.

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

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

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

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

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

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

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

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



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

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

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

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

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

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

OPSAPS-68148: Cruise Control rack aware goal upgrade handler

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

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

Known Issues in Data Analytics Studio

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

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

Workaround: To obtain the DAS WebApp logs, check the `stderr.log` file in the runtime process directory for the DAS WebApp.

- You may not be able to add or delete columns or change the table schema after creating a new table using the upload table feature.
- For clusters secured using Knox, you see the HTTP 401: Forbidden error message when you click the DAS quick link from Cloudera Manager and are unable to log into DAS.

Workaround: The admin user will need to provide the DAS URL from the Knox proxy topology to the users needing access to DAS.

- The download logs feature may not return the YARN application logs on a Kerberized cluster. When you download the logs, the logs contain an `error-reports.json` file which states that no valid Kerberos tokens are available.

Workaround: An admin user with access to the machine can use the `kinit` command as a hive user with hive service user keytabs and trigger the download.

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

Workaround:

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

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

3. Start the DAS Event Processor.

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

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

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

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

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

```
select * from test where option = '--name';
```

- You might face UI issues on Google Chrome while using faceted search. We recommend you to use the latest version of Google Chrome (version 71.x or higher).
- Visual Explain for the same query shows different graphs on the **Compose** page and the **Query Details** page.

- While running some queries, if you restart HSI, the query execution is stopped. However, DAS does not reflect this change and the queries appear to be in the same state forever.
- After a fresh installation, when there is no data and you try to access the Reports tab, DAS displays an "HTTP 404 Not Found" error.
- Join count does not get updated for tables with partitioned columns.

Known Issues in Apache HBase

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

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

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

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

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

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

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

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

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

Apache Issue: N/A

Known Issues in HDFS

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

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

None.

Unsupported Features

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

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

Technical Service Bulletins

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

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

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

Knowledge article

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

Known Issues in Apache Hive

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

CDPD-66549: Cloudera Machine Learning failed to connect to Hive Metastore during an upgrade

Hive Metastore (HMS) becomes unreachable during and after an upgrade from Cloudera Runtime 7.2.17 to higher versions.

Clients communicate with HMS using delegation tokens and these tokens that are required during a connection, is stored in memory. After an upgrade, all issued tokens are lost when the HMS restarts, and as a result the handshake communication using the old token is unsuccessful. If you set the HMS Delegation Token Store to DBTokenStore, HMS will persist the tokens in the backend database. The tokens can then be retrieved after an upgrade since it is stored ensuring that the handshake communication using the old token is successful.

Perform the following steps to configure the HMS Delegation Token Store:

1. In Cloudera Manager, click Clusters HIVE Configuration
2. Search for the hive.cluster.delegation.token.store.class property and set the value to org.apache.hadoop.hive.thrift.DBTokenStore
3. Restart the Hive service for the changes to take effect.

CDPD-54988: Disallow creation of Temporary Hive Iceberg tables

Atlas by default skips ingesting temporary tables created in Hive. But when searched for the newly created temporary Hive tables, it displays them in the search results.

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

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

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

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

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

Update Hive log4j configurations. Hive -> Configuration -> HiveServer2 Logging Advanced Configuration Snippet (Safety Valve) Hive Metastore -> Configuration -> Hive Metastore Server Logging Advanced Configuration Snippet (Safety Valve) Add

the following to the configurations: appender.DRFA.strategy.action.type=DELETE
 appender.DRFA.strategy.action.basepath=\${log.dir} appender.DRFA.strategy.action.maxdepth=1
 appender.DRFA.strategy.action.PathConditions.glob=\${log.file}.*
 appender.DRFA.strategy.action.PathConditions.type=IfFileName
 appender.DRFA.strategy.action.PathConditions.nestedConditions.type=IfAccumulatedFileCount
 appender.DRFA.strategy.action.PathConditions.nestedConditions.exceeds=same value as
 appender.DRFA.strategy.max

CDPD-10848: HiveServer Web UI displays incorrect data

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

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

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

Use one of the workarounds:

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

Known Issues in Hue

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

Known issues in 7.2.17

CDPD-56888: Renaming a folder with special characters results in a duplicate folder with a new name on AWS S3.

On AWS S3, if you try to rename a folder with special characters in its name, a new folder is created as a copy of the original folder with its contents. Also, you may not be able to delete the folder containing special characters.

You can rename or delete a directory having special characters using the HDFS commands as follows:

1. SSH into your CDP environment host.
2. To delete a directory within your S3 bucket, run the following command:

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

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

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

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

CDPD-48146: Error while browsing S3 buckets or ADLS containers from the left-assist panel

You may see the following error while trying to access the S3 buckets or ADLS containers from the left-assist panel in Hue: Failed to retrieve buckets: :1:0: syntax error.

Access the S3 buckets or ADLS containers using the File Browser.

CDPD-54376: Clicking the home button on the File Browser page redirects to HDFS user directory

When you are previewing a file on any supported filesystem, such as S3 or ABFS, and you click on the Home button, you are redirected to the HDFS user home directory instead of the user home directory on the said filesystem.

None.

Technical Service Bulletins

TSB 2023-704: File corruption when downloading files larger than 1 MB from ABFS with Hue File Browser

An issue within the upstream Apache Hue (Hue) application results in file corruption when downloading files larger than 1MB files from Azure Blob Filesystem (ABFS) using the Hue File Browser. Only the downloaded files are affected by this issue, the source files remain intact.

Knowledge article

For the latest update on this issue, see the corresponding Knowledge article: [TSB 2023-704: File corruption when downloading files larger than 1 MB from ABFS with Hue File Browser](#)

TSB 2023-703: Risk of Data Loss when using Hue S3 File Browser

There is a risk of losing data when moving one or more files or folders in Amazon S3 storage with the Hue File Browser. When the user selects the destination folder in the modal window the following scenarios can occur:

1. If the user selects the same destination folder as the source folder, and clicks the Move button, the selected files will be permanently deleted.
2. If the user selects a different destination folder from the source folder and clicks the Move button before the User Interface (UI) has completely loaded (the loading is indicated by a spinner), the action could lead to the following outcomes:
 - a. If the previously visible destination folder was the same as the source folder, the file will be permanently deleted.
 - b. If the previously visible destination folder was different from the source folder, the file(s) will be moved to the previously visible destination folder and not to the intended destination folder.

Knowledge article

For the latest update on this issue, see the corresponding Knowledge article: [TSB 2023-703: Risk of Data Loss when using Hue S3 File Browser](#)

Known issues in 7.2.16

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

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

See [Authenticating Hue users with Knox SSO](#).

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

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

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

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

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

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

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

CDPD-43293: Unable to import Impala table using Importer

Creating Impala tables using the Hue Importer may fail.

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

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

Known issues before 7.2.16

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

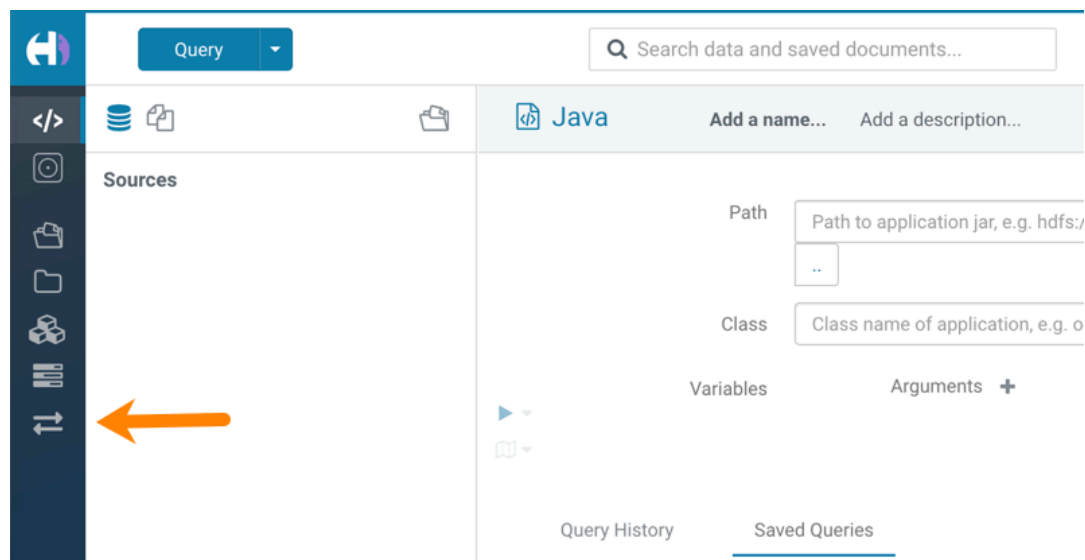
When you run Impala queries in a batch mode, you encounter failures with a Kerberos error even if the keytab is configured correctly. This is because submitting Impala, Sqoop, Pig, or pyspark queries in a batch mode launches a shell script Oozie job from Hue and this is not supported on a secure cluster.

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

Hue Importer is not supported in the Data Engineering template

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

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



Unsupported features

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

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

```
[desktop]
[spark]
```

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

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

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

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

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

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



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

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

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

None.

Known Issues Iceberg

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

Concurrent compactions and modify statements can corrupt Iceberg tables

Hive or Impala DELETE/UPDATE/MERGE operations on Iceberg V2 tables can corrupt the tables if there is a concurrent table compaction from Spark. The issue happens if the compaction and modify statement runs in parallel, and if the compaction job commits before the modify statement. In that case the modify statement's position delete files still point to the old files. The results in the case of DELETE and in the case of UPDATE / MERGE are as follows:

- DELETE
 - Delete records pointing to old files have no effect.
- UPDATE / MERGE
 - Delete records pointing to old files have no effect. The table will also have the newly added data records, which means rewritten records will still be active.

Use one of the following workarounds:

- Do not run compactions and DELETE/UPDATE/MERGE statements in parallel.
- Do not compact the table via Iceberg's RewriteFiles operation. For example do not use Spark's `rewriteDataFiles`.

CDPD-57551: Performance issue can occur on reads after writes of Iceberg tables

Hive might generate too many small files, which causes performance degradation.

Maintain a relatively small number of data files under the iceberg table/partition directory to have efficient reads. To alleviate poor performance caused by too many small files, run the following queries:

```
TRUNCATE TABLE target;
INSERT OVERWRITE TABLE target select * from target FOR SYSTEM_VERSION AS OF <preTruncateSnapshotId>;
```

Known Issues in Apache Impala

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

CDPD-41138: Reading through <https://github.com/hunterhacker/jdom/issues/189>, the fix for CVE-2021-33813 is specifically that if you were relying on `setFeature("http://xml.org/sax/features/external-general-entities", false)`, it was not applied correctly and you were still vulnerable. However if you used `setExpandEntities(false)` then you're not vulnerable to CVE-2021-33813.

I found sources for rome 0.9 at <http://www.java2s.com/Code/Jar/t/Downloadrome09sourcesjar.htm> (it's no longer available at <https://java.net/>) and verified it uses both `setFeature` and `setExpandEntities` to prevent XXE attacks. So I don't believe rome in particular is vulnerable to this issue, and jdom 1.0 is only included for rome 0.9.

None

Impala known limitation when querying compacted tables

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

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

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

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`

This vulnerability is fixed upstream under [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](#)

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

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

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

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

Restart the impalad.

IMPALA-532: Impala should tolerate bad locale settings

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

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

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

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

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

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

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

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

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

IMPALA-635: Avro Scanner fails to parse some schemas

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

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

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

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

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

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

Impala is not producing these logs.

None

IMPALA-1024: Impala BE cannot parse Avro schema that contains a trailing semi-colon

If an Avro table has a schema definition with a trailing semicolon, Impala encounters an error when the table is queried.

Remove trailing semicolon from the Avro schema.

IMPALA-1652: Incorrect results with basic predicate on CHAR typed column

When comparing a CHAR column value to a string literal, the literal value is not blank-padded and so the comparison might fail when it should match.

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

IMPALA-1792: ImpalaODBC: Can not get the value in the SQLGetData(m-x th column) after the SQLBindCol(m th column)

If the ODBC `SQLGetData` is called on a series of columns, the function calls must follow the same order as the columns. For example, if data is fetched from column 2 then column 1, the `SQLGetData` call for column 1 returns NULL.

Fetch columns in the same order they are defined in the table.

IMPALA-1821: Casting scenarios with invalid/inconsistent results

Using a `CAST()` function to convert large literal values to smaller types, or to convert special values such as NaN or Inf, produces values not consistent with other database systems. This could lead to unexpected results from queries.

IMPALA-2005: A failed CTAS does not drop the table if the insert fails

If a `CREATE TABLE AS SELECT` operation successfully creates the target table but an error occurs while querying the source table or copying the data, the new table is left behind rather than being dropped.

Drop the new table manually after a failed `CREATE TABLE AS SELECT`

IMPALA-2422: % escaping does not work correctly when occurs at the end in a LIKE clause

If the final character in the RHS argument of a `LIKE` operator is an escaped `\%` character, it does not match a `%` final character of the LHS argument.

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

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

IMPALA-3094: Incorrect result due to constant evaluation in query with outer join

An `OUTER JOIN` query could omit some expected result rows due to a constant such as `FALSE` in another join clause. For example:

```
explain SELECT 1 FROM alltypestiny a1
INNER JOIN alltypesagg a2 ON a1.smallint_col = a2.year AND fals
e
RIGHT JOIN alltypes a3 ON a1.year = a1.bigint_col;
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--+

```


it as the final output. However, in cases where there are still results to return after this timeout has passed, when the driver closes the connection, based on the timeout, it results in a scenario where the query results are incomplete.

Upstream JIRA

[IMPALA-7561](#)

Knowledge article

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

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

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

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

Upstream JIRA

[IMPALA-11030](#)

Knowledge article

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

Known Issues in Apache Kafka

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

Known Issues

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

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

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

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

The `offsets.topic.replication.factor` property must be less than or equal to the number of live brokers

The `offsets.topic.replication.factor` broker configuration is now enforced upon auto topic creation. Internal auto topic creation will fail with a `GROUP_COORDINATOR_NOT_AVAILABLE` error until the cluster size meets this replication factor requirement.

None

Requests fail when sending to a nonexistent topic with `auto.create.topics.enable` set to true

The first few produce requests fail when sending to a nonexistent topic with `auto.create.topics.enable` set to true.

Increase the number of retries in the producer configuration setting `retries`.

KAFKA-2561: Performance degradation when SSL is enabled

In some configuration scenarios, significant performance degradation can occur when SSL is enabled. The impact varies depending on your CPU, JVM version, Kafka configuration, and message size. Consumers are typically more affected than producers.

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

OPSAPS-43236: Kafka garbage collection logs are written to the process directory

By default Kafka garbage collection logs are written to the agent process directory. Changing the default path for these log files is currently unsupported.

None

CDPD-45183: Kafka Connect active topics might be visible to unauthorised users

The Kafka Connect active topics endpoint (`/connectors/[***CONNECTOR NAME***/topics`) and the Connect Cluster page on the SMM UI disregard the user permissions configured for the Kafka service in Ranger. As a result, all active topics of connectors might become visible to users who do not have permissions to view them. Note that user permission configured for Kafka Connect in Ranger are not affected by this issue and are correctly applied.

None.

RANGER-3809: Idempotent Kafka producer fails to initialize due to an authorization failure

Kafka producers that have idempotence enabled require the Idempotent Write permission to be set on the cluster resource in Ranger. If permission is not given, the client fails to initialize and an error similar to the following is thrown:

```
org.apache.kafka.common.KafkaException: Cannot execute transactional method because we are in an error state
    at org.apache.kafka.clients.producer.internals.TransactionManager.maybeFailWithError(TransactionManager.java:1125)
    at org.apache.kafka.clients.producer.internals.TransactionManager.maybeAddPartition(TransactionManager.java:442)
    at org.apache.kafka.clients.producer.KafkaProducer.doSend(KafkaProducer.java:1000)
    at org.apache.kafka.clients.producer.KafkaProducer.send(KafkaProducer.java:914)
    at org.apache.kafka.clients.producer.KafkaProducer.send(KafkaProducer.java:800)
    .
    .
    .
Caused by: org.apache.kafka.common.errors.ClusterAuthorizationException: Cluster authorization failed.
```

Idempotence is enabled by default for clients in Kafka 3.0.1, 3.1.1, and any version after 3.1.1. This means that any client updated to 3.0.1, 3.1.1, or any version after 3.1.1 is affected by this issue.

This issue has two workarounds, do either of the following:

- Explicitly disable idempotence for the producers. This can be done by setting `enable.idempotence` to `false`.
- Update your policies in Ranger and ensure that producers have Idempotent Write permission on the cluster resource.

DBZ-4990: The Debezium Db2 Source connector does not support schema evolution

The Debezium Db2 Source connector does not support the evolution (updates) of schemas. In addition, schema change events are not emitted to the schema change topic if there is a change in the schema of a table that is in capture mode. For more information, see [DBZ-4990](#).

None.

Unsupported Features

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

- Only Java and .Net based clients are supported. Clients developed with C, C++, Python, and other languages are currently not supported.
- The Kafka default authorizer is not supported. This includes setting ACLs and all related APIs, broker functionality, and command-line tools.
- SASL/SCRAM is only supported for delegation token based authentication. It is not supported as a standalone authentication mechanism.
- Kafka KRaft in this release of Cloudera Runtime is in technical preview and does not support the following:
 - Deployments with multiple log directories. This includes deployments that use JBOD for storage.
 - Delegation token based authentication.
 - Migrating an already running Kafka service from ZooKeeper to KRaft.
 - Atlas Integration.

Limitations

Collection of Partition Level Metrics May Cause Cloudera Manager's Performance to Degrade

If the Kafka service operates with a large number of partitions, collection of partition level metrics may cause Cloudera Manager's performance to degrade.

If you are observing performance degradation and your cluster is operating with a high number of partitions, you can choose to disable the collection of partition level metrics.



Important: If you are using SMM to monitor Kafka or Cruise Control for rebalancing Kafka partitions, be aware that both SMM and Cruise Control rely on partition level metrics. If partition level metric collection is disabled, SMM will not be able to display information about partitions. In addition, Cruise Control will not operate properly.

Complete the following steps to turn off the collection of partition level metrics:

1. Obtain the Kafka service name:
 - a. In Cloudera Manager, Select the Kafka service.
 - b. Select any available chart, and select Open in Chart Builder from the configuration icon drop-down.
 - c. Find \$SERVICENAME= near the top of the display.
The Kafka service name is the value of \$SERVICENAME.
2. Turn off the collection of partition level metrics:
 - a. Go to Hosts Configuration.
 - b. Find and configure the Cloudera Manager Agent Monitoring Advanced Configuration Snippet (Safety Valve) configuration property.

Enter the following to turn off the collection of partition level metrics:

```
[KAFKA_SERVICE_NAME]_feature_send_broker_topic_partition_entity_update_enabled=false
```

Replace [KAFKA_SERVICE_NAME] with the service name of Kafka obtained in step 1. The service name should always be in lower case.

- c. Click Save Changes.

Known Issues in Apache Knox

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

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

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

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

CDPD-60376: Cloud loadbalancer takes 20-30 secs to failover to the next available Knox host

If Knox is in HA and one of the Knox server is down, then accessing of service via Control plane endpoint url(i.e. via cloud loadbalancer) will take ~ 30secs to failover the request to available Knox instance .

Retry the request after 30 seconds.

Known Issues in Apache Kudu

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

Kudu supports both coarse-grain and fine-grain authorization, but Kudu does not yet support integration with Atlas.

None

CDPD-57181: The kudu service user is not authorized to access Hive warehouse locations on cloud object stores which can prevent Kudu tables to be created under certain conditions..

Add "kudu" to the allow list for "Default: Hive warehouse locations" in the Ranger repository for your object storage.

Known Issues in Apache Oozie

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

CDPD-41274: HWC + Oozie issue: Could not open client transport with JDBC Uri

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

Use Spark action in cluster mode.

```
Use Spark action in cluster mode.
    <spark xmlns="uri:oozie:spark-action:1
.0">
    ...
    <mode>cluster</mode>
    ...
</spark>
```

Oozie jobs fail (gracefully) on secure YARN clusters when JobHistory server is down

If the JobHistory server is down on a YARN (MRv2) cluster, Oozie attempts to submit a job, by default, three times. If the job fails, Oozie automatically puts the workflow in a SUSPEND state.

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

Unsupported Feature

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

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

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

Known Issues in Apache Phoenix

There are no known issues for Phoenix in Cloudera Runtime 7.2.17.

Known Issues in Apache Ranger

Learn about the known issues in Apache Ranger, the impact or changes to the functionality, and the workaround.
OPSAPS-70387: The DataHub cluster deletion process does not delete the Ranger entries which created for the same cluster.

If the user wants to create a new DataHub cluster with same old name then it fails because as there was an entry with the same name already in Ranger.

User must delete the Ranger entries manually which contains the DataHub cluster name.

Known Issues in Schema Registry

Learn about the known issues in Schema Registry, the impact or changes to the functionality, and the workaround.
CDPD-56890: New schemas cannot be created following an upgrade

If you delete the latest version of a schema (the one with the highest ID) from the Schema Registry database before an upgrade, you might not be able to create new schemas after you upgrade the cluster to a newer version.



Important: In CDP Public Cloud, this issue only manifests when upgrading from Cloudera Runtime 7.2.12 or lower to 7.2.14 or higher.

1. Access the Schema Registry database. Go to Cloudera Manager Schema Registry Configuration and search for "database" if you don't know the name, host, or port of the Schema Registry database.
2. Cross reference the ID's in the schemaVersionId column of the schmema_version_state table with the ID's found in the schema_version_info table.
3. Delete all records from the schema_version_state table that contains a schemaVersionId not present in the schema_version_info table.

CDPD-58265: Schema Registry Client incorrectly applies SSL configuration

The Cloudera distributed Schema Registry Java client might fail to apply the SSL configurations correctly with concurrent access in Jersey clients due to a [Jersey](#) issue related to JDK.

Before using `HttpsURLConnection` in any form concurrently, call `javax.net.ssl.HttpsURLConnection.getDefaultSSLContextFactory()` once in the custom client application.

CDPD-58949: Schemas are de-duplicated on import

On import, Schema Registry de-duplicates schema versions based on their fingerprints. This means that schemas which are considered functionally equivalent in SR get de-duplicated. As a result, some schema versions are not created, and their IDs do not become valid IDs in SR.

None.

CDPD-55381: Schema Registry issues authentication cookie for the authorized user, not for the authenticated one

When the authenticated user is different from the authorized user, which can happen when Schema Registry is used behind Knox, authorization issues can occur for subsequent requests as the authentication cookie in Schema Registry stores the authorized user.

Access Schema Registry directly, without using Knox, if possible. If not, ensure that the name of the end user that tries to connect does not begin with knox.

CDPD-60160: Schema Registry Atlas integration does not work with Oracle databases

Schema Registry is unable to create entities in Atlas if Schema Registry uses an Oracle database. The following will be present in the Schema Registry log if you are affected by this issue:

```
ERROR com.cloudera.dim.atlas.events.AtlasEventsProcessor: An error occurred while processing Atlas events.
java.lang.IllegalArgumentException: Cannot invoke com.hortonworks.registries.schemaregistry.AtlasEventStorable.setType on bean class 'class com.hortonworks.registries.schemaregistry.AtlasEventStorable' - argument type mismatch - had objects of type "java.lang.Long" but expected signature "java.lang.Integer"
```

This issue causes the loss of audit data on Oracle environments.

None.

CDPD-59015: Schema Registry does not create new versions of schemas even if the schema is changed

Schema Registry uses a schema fingerprinting mechanism to differentiate between schemas. However, fingerprinting does not take into consideration the schema attributes of the field type. As a result, if you have two schemas where the only difference is that one has type attributes defined and the other does not, they will be considered identical by Schema Registry. For example, the following schemas are considered identical:

```
#Schema V1
{"type": "record", "name": "schema_name", "namespace": "ns", "fields": [
  {"name": "local_timestamp_micros_long", "type": "long"}]}

#Schema V2
{"type": "record", "name": "schema_name", "namespace": "ns", "fields": [
  {"name": "local_timestamp_micros_long", "type": {"type": "long", "logicalType": "local-timestamp-micros"}}]}
```

Notice that the only difference is that in the second schema, the `local_timestamp_micros_long` field has a logical type specified. In cases like this, the new version of the schema is not created, the initial version is used. This is true even if the data that is being produced has a new schema version. The ID of the first schema version is used and is put in the serialized record. The new schema version is not created.

This issue is common when using change data capture (CDC) connectors like the Debezium connectors. This is because CDC connectors create schemas with the logical type decimal based on the column type in the database schema. For example:

```
{"type": "record", "name": "schema_name", "namespace": "ns", "fields": [
  {"name": "database_column", "type": {"type": "bytes", "logicalType": "decimal", "precision": 64, "scale": 0}}]}
```

If the database schema changes (for example, the column type), it is possible that only scale changes, which is a schema attribute.

```
{"type": "record", "name": "schema_name", "namespace": "ns", "fields": [
  {"name": "database_column", "type": {"type": "bytes", "logicalType": "decimal", "precision": 64, "scale": 1}}]}
```

In this case, even though scale changed to 1, the first version of the schema is used where scale is 0. As a result, the data is consumed with the wrong scale.

Avoid using logical types or other attributes. Alternatively, ensure that there are no changes in the logical types or other attributes between schema versions.

OPSAPS-68708: Schema Registry might fail to start if a load balancer address is specified in Ranger

Schema Registry does not start if the address specified in the Load Balancer Address Ranger property does not end with a trailing slash (/).

Set the value of the RANGER_REST_URL Schema Registry environment variable to an address that includes a trailing slash.

1. In Cloudera Manager, select the Schema Registry service.
2. Go to Configuration.
3. Find the Schema Registry Server Environment Advanced Configuration Snippet (Safety Valve) property and add the following:

```
Key: RANGER_REST_URL
Value: [***RANGER REST API URL***]
```

Replace [***RANGER REST API URL***] with an address that can be used by Schema Registry to access Ranger. Ensure that the address ends with a trailing slash. For example: `http://ranger-1.cloudera.com:6182/`

4. Restart the Schema Registry service.

CDPD-58990: getSortedSchemaVersions method orders by schemaVersionId instead of version number

On validation, Schema Registry orders schema versions based on ID instead of version number. In some situations, this can cause validation with the LATEST level to compare the new schema version to a non-latest version.

This situation can occur when an older version of a schema has a higher ID than the newer version of a schema, for example, when the older version is imported with an explicit ID.

None.

Known Issues in Apache Solr

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

Known Issues

Changing the default value of Client Connection Registry HBase configuration parameter causes HBase MRIT job to fail

If the value of the HBase configuration property Client Connection Registry is changed from the default ZooKeeper Quorum to Master Registry then the Yarn job started by HBase MRIT fails with a similar error message:

```
Caused by: org.apache.hadoop.hbase.exceptions.MasterRegistryFetchException: Exception making rpc to masters [quasar-bmyccr-2.quasar-bmyccr.root.hwx.site,22001,-1]
    at org.apache.hadoop.hbase.client.MasterRegistry.lambda$groupCall$1(MasterRegistry.java:244)
    at org.apache.hadoop.hbase.util.FutureUtils.lambda$addListener$0(FutureUtils.java:68)
    at java.util.concurrent.CompletableFuture.uniWhenComplete(CompletableFuture.java:774)
    at java.util.concurrent.CompletableFuture.uniWhenCompleteStage(CompletableFuture.java:792)
    at java.util.concurrent.CompletableFuture.whenComplete(CompletableFuture.java:2153)
    at org.apache.hadoop.hbase.util.FutureUtils.addListener(FutureUtils.java:61)
```

```

    at org.apache.hadoop.hbase.client.MasterRegistry.groupCall(MasterRegistry.java:228)
    at org.apache.hadoop.hbase.client.MasterRegistry.call(MasterRegistry.java:265)
    at org.apache.hadoop.hbase.client.MasterRegistry.getMetaRegionLocations(MasterRegistry.java:282)
    at org.apache.hadoop.hbase.client.ConnectionImplementation.locateMeta(ConnectionImplementation.java:900)
    at org.apache.hadoop.hbase.client.ConnectionImplementation.locateRegion(ConnectionImplementation.java:867)
    at org.apache.hadoop.hbase.client.ConnectionImplementation.relocateRegion(ConnectionImplementation.java:850)
    at org.apache.hadoop.hbase.client.ConnectionImplementation.locateRegionInMeta(ConnectionImplementation.java:981)
    at org.apache.hadoop.hbase.client.ConnectionImplementation.locateRegion(ConnectionImplementation.java:870)
    at org.apache.hadoop.hbase.client.RpcRetryingCallerWithReadReplicas.getRegionLocations(RpcRetryingCallerWithReadReplicas.java:319)
    ... 21 more
Caused by: org.apache.hadoop.hbase.client.RetriesExhaustedException: Failed contacting masters after 1 attempts.
Exceptions:
java.io.IOException: Call to address=quasar-bmyccr-2.quasar-bmyccr.root.hwx.site/172.27.19.4:22001 failed on local exception: java.io.IOException: java.lang.RuntimeException: Found no valid authentication method from options
    at org.apache.hadoop.hbase.client.MasterRegistry.lambda$groupCall$1(MasterRegistry.java:243)
    ... 35 more

```

Add the following line to the MRIT command line:

```
-D 'hbase.client.registry.impl=org.apache.hadoop.hbase.client.ZKConnectionRegistry'
```

Solr does not support rolling upgrade to release 7.2.18 or lower

Solr supports rolling upgrades from release 7.2.18 and higher. Upgrading from a lower version means that all the Solr Server instances are shut down, parcels upgraded and activated and then the Solr Servers are started again. This causes a service interruption of several minutes, the actual value depending on cluster size.

Services like Atlas and Ranger that depend on Solr, may face issues because of this service interruption.

None.

Cannot create multiple heap dump files because of file name error

Heap dump generation fails with a similar error message:

```

java.lang.OutOfMemoryError: Java heap space
Dumping heap to /data/tmp/solr_solr-SOLR_SERVER-fc9dacc265fabfc500b92112712505e3_pid{{PID}}.hprof ...
Unable to create /data/tmp/solr_solr-SOLR_SERVER-fc9dacc265fabfc500b92112712505e3_pid{{PID}}.hprof: File exists

```

The cause of the problem is that {{PID}} does not get substituted during dump file creation with an actual process ID and because of that, a generic file name is generated. This causes the next dump file creation to fail, as the existing file with the same name cannot be overwritten.

You need to manually delete the existing dump file.

Solr coreAdmin status throws Null Pointer Exception

You get a Null Pointer Exception with a similar stacktrace:

```
Caused by: java.lang.NullPointerException
    at org.apache.solr.core.SolrCore.getInstancePath(SolrCore.java:333)
    at org.apache.solr.handler.admin.CoreAdminOperation.getCoreStatus(CoreAdminOperation.java:324)
    at org.apache.solr.handler.admin.StatusOp.execute(StatusOp.java:46)
    at org.apache.solr.handler.admin.CoreAdminOperation.execute(CoreAdminOperation.java:362)
```

This is caused by an error in handling solr admin core STATUS after collections are rebuilt.

Restart the Solr server.

Applications fail because of mixed authentication methods within dependency chain of services

Using different types of authentication methods within a dependency chain, for example, configuring your indexer tool to authenticate using Kerberos and configuring your Solr Server to use LDAP for authentication may cause your application to time out and eventually fail.

Make sure that all services in a dependency chain use the same type of authentication.

API calls fail with error when used with alias, but work with collection name

API calls fail with a similar error message when used with an alias, but they work when made using the collection name:

```
[ ] o.a.h.s.t.d.w.DelegationTokenAuthenticationFilter Authentication exception: User: xyz@something.example.com is not allowed to impersonate xyz@something.example.com
[c:RTOTagMetaOdd s:shard3 r:core_node11 x:RTOTagMetaOdd_shard3_replica_n8] o.a.h.s.t.d.w.DelegationTokenAuthenticationFilter Authentication exception: User: xyz@something.example.com is not allowed to impersonate xyz@something.example.com
```

Make sure there is a replica of the collection on every host.

Apache Tika upgrade may break morphlines indexing

The upgrade of Apache Tika from 1.27 to 2.3.0 brought potentially breaking changes for morphlines indexing. Duplicate/triplicate keys names were removed and certain parser class names were changed (For example, org.apache.tika.parser.jpeg.JpegParser changed to org.apache.tika.parser.image.JpegParser).

To avoid morphline commands failing after the upgrade, do the following:

- Check if key name changes affect your morphlines. For more information, see *Removed duplicate/triplicate keys* in [Migrating to Tika 2.0.0](#).
- Check if the name of any parser you use has changed. For more information, see the Apache Tika [API documentation](#).

Update your morphlines if necessary.

CDPD-28432: HBase Lily indexer REST port does not support SSL

When using the --http argument for the hbase-indexer command line tool to invoke Lily indexer through REST API, you can add/list/remove indexers with any user without the need for authentication. Keeping the default true value for the hbaseindexer.httpservers.disabled environment parameter switches off the REST interface, so no one can use the --http argument when using the hbase-indexer command line tool. This also means that users need to authenticate as an hbase user in order to use the hbase-indexer tool.

CDH-77598: Indexing fails with socketTimeout

Starting from CDH 6.0, the HTTP client library used by Solr has a default socket timeout of 10 minutes. Because of this, if a single request sent from an indexer executor to Solr takes more than 10 minutes to be serviced, the indexing process fails with a timeout error.

This timeout has been raised to 24 hours. Nevertheless, there still may be use cases where even this extended timeout period proves insufficient.

If your MapreduceIndexerTool or HBaseMapreduceIndexerTool batch indexing jobs fail with a timeout error during the go-live (Live merge, MERGEINDEXES) phase (This means the merge takes longer than 24 hours).

Use the `--go-live-timeout` option where the timeout can be specified in milliseconds.

CDPD-12450: CrunchIndexerTool Indexing fails with socketTimeout

The http client library uses a socket timeout of 10 minutes. The Spark Crunch Indexer does not override this value, and in case a single batch takes more than 10 minutes, the entire indexing job fails. This can happen especially if the morphlines contain DeleteByQuery requests.

Try the following workarounds:

- Check the batch size of your indexing job. Sending too large batches to Solr might increase the time needed on the Solr server to process the incoming batch.
- If your indexing job uses `deleteByQuery` requests, consider using `deleteById` wherever possible as `deleteByQuery` involves a complex locking mechanism on the Solr side which makes processing the requests slower.
- Check the number of executors for your Spark Crunch Indexer job. Too many executors can overload the Solr service. You can configure the number of executors by using the `--mappers` parameter
- Check that your Solr installation is correctly sized to accommodate the indexing load, making sure that the number of Solr servers and the number of shards in your target collection are adequate.
- The socket timeout for the connection can be configured in the morphline file. Add the `solrClientSocketTimeout` parameter to the `solrLocator` command

Example

```
SOLR_LOCATOR :
{
  collection : test_collection
  zkHost : "zookeeper1.example.corp:2181/solr"
  # 10 minutes in milliseconds
  solrClientSocketTimeout: 600000
  # Max number of documents to pass per RPC from morphline to
  Solr Server
  # batchSize : 10000
}
```

CDPD-29289: HBaseMapReduceIndexerTool fails with socketTimeout

The http client library uses a socket timeout of 10 minutes. The HBase Indexer does not override this value, and in case a single batch takes more than 10 minutes, the entire indexing job fails.

You can overwrite the default 600000 millisecond (10 minute) socket timeout in HBase indexer using the `--solr-client-socket-timeout` optional argument for the direct writing mode (when the value of the `--reducers` optional argument is set to 0 and mappers directly send the data to the live Solr).

CDPD-20577: Splitshard operation on HDFS index checks local filesystem and fails

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

Run the following command to skip the check for sufficient disk space altogether:

- On nonsecure clusters:

```
curl 'http://[${**SOLR_SERVER_HOSTNAME**}]:8983/solr/admin/collections?action=SPLITSHARD&collection=[**COLLECTION_NAME**]&shard=[**SHARD_TO_SPLIT**]&skipFreeSpaceCheck=true'
```

- On secure clusters:

```
curl -k -u : --negotiate 'http://[${**SOLR_SERVER_HOSTNAME**}]:8985/solr/admin/collections?action=SPLITSHARD&collection=[**COLLECTION_NAME**]&shard=[**SHARD_TO_SPLIT**]&skipFreeSpaceCheck=true'
```

Replace `[**SOLR_SERVER_HOSTNAME**]` with a valid Solr server hostname, `[**COLLECTION_NAME**]` with the collection name, and `[**SHARD_TO_SPLIT**]` with the ID of the shard to split.

To verify that the command executed successfully, check overseer logs for a similar entry:

```
2021-02-02 12:43:23.743 INFO (OverseerThreadFactory-9-thread-5-processing-n:myhost.example.com:8983_solr) [c:example s:shard1] o.a.s.c.a.c.SplitShardCmd Skipping check for sufficient disk space
```

DOCS-5717: Lucene index handling limitation

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

There is no workaround, you need to reindex collections.

CDH-22190: CrunchIndexerTool which includes Spark indexer requires specific input file format specifications

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

None

CDH-26856: Field value class guessing and Automatic schema field addition are not supported with the MapReduceIndexerTool nor with the HBaseMapReduceIndexerTool.

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

Define the schema before running the `MapReduceIndexerTool` or `HBaseMapReduceIndexerTool`. In non-schemaless mode, define in the schema using the `schema.xml` file. In schemaless mode, either define the schema using the Solr Schema API or index sample documents using NRT indexing before invoking the tools. In either case, Cloudera recommends that you verify that the schema is what you expect, using the List Fields API command.

CDH-19407: The Browse and Spell Request Handlers are not enabled in schemaless mode

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

If you require the Browse and Spell Request Handlers, add them to the `solrconfig.xml` configuration file. Generate a non-schemaless configuration to see the usual settings and modify the required fields to fit your schema.

CDH-17978: Enabling blockcache writing may result in unusable indexes.

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

None

CDH-58276: Users with insufficient Solr permissions may receive a "Page Loading" message from the Solr Web Admin UI.

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

None

CDH-15441: Using MapReduceIndexerTool or HBaseMapReduceIndexerTool multiple times may produce duplicate entries in a collection.

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

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



Note: This workaround is only valid for HBaseMapReduceIndexerTool. There is no workaround for MapReduceIndexerTool

CDH-58694: Deleting collections might fail if hosts are unavailable.

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

Ensure all hosts are online before deleting collections.

Unsupported Features

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

- Package Management System
- HTTP/2
- Solr SQL/JDBC
- Graph Traversal
- Cross Data Center Replication (CDCR)
- SolrCloud Autoscaling
- HDFS Federation
- Saving search results
- Solr contrib modules (Spark, MapReduce and Lily HBase indexers are not contrib modules but part of Cloudera's distribution of Solr itself, therefore they are supported).

Known Issues in Apache Spark

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

CDPD-217: The Apache Spark connector is not supported

The old *Apache Spark - Apache HBase Connector* (shc) is not supported in CDP releases.

Use the new HBase-Spark connector shipped in CDP release.

CDPD-3038: Launching pyspark displays several HiveConf warning messages

When pyspark starts, several Hive configuration warning messages are displayed, similar to the following:

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

These errors can be safely ignored.

Known Issues in Apache Spark3

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

CDPD-57989: MERGE INTO Query fails on tables with non-nullable columns.

None

Known Issues for Apache Sqoop

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

Using direct mode causes problems

Using direct mode has several drawbacks:

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

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

Sqoop's direct mode is no longer supported and is disabled by default. However, if you still want to use it, enable it by either setting the `sqoop.enable.deprecated.direct` property globally in Cloudera Manager for Sqoop or by specifying it in the command-line through `-Dsqoop.enable.deprecated.direct=true`.

CDPD-3089: Avro, S3, and HCat do not work together properly

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

Parquet columns inadvertently renamed

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

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

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

Importing multiple megabytes per row before initial-page-run check (ColumnWriter) can cause OOM. Also, rows that vary significantly by size so that the next-page-size check is based on small rows, and is set very high, followed by many large rows can also cause OOM.

None

Known issues in Streams Messaging Manager

Learn about the known issues in Streams Messaging Manager, the impact or changes to the functionality, and the workaround.

CDPD-39313: Some numbers are not rendered properly in SMM UI

Very large numbers can be imprecisely represented on the UI. For example, bytes larger than 8 petabytes would lose precision.

None.

CDPD-45183: Kafka Connect active topics might be visible to unauthorised users

The Kafka Connect active topics endpoint (/connectors/[***CONNECTOR NAME***/topics) and the Connect Cluster page on the SMM UI disregard the user permissions configured for the Kafka service in Ranger. As a result, all active topics of connectors might become visible to users who do not have permissions to view them. Note that user permission configured for Kafka Connect in Ranger are not affected by this issue and are correctly applied.

None.

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

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

SMM cannot identify multiple listeners and still points to bootstrap server using the default broker port (9093 for SASL_SSL). You would have to override bootstrap server URL (hostname:port as set in the listeners for broker). Add the bootstrap server details in SMM safety valve in the following path:

Cloudera Manager SMM Configuration Streams Messaging Manager Rest Admin Server Advanced Configuration Snippet (Safety Valve) for streams-messaging-manager.yaml Add the following value for bootstrap servers Save Changes Restart SMM .

```
streams.messaging.manager.kafka.bootstrap.servers=<comma-separated list of brokers>
```

OPSAPS-59597: SMM UI logs are not supported by Cloudera Manager

Cloudera Manager does not support the log type used by SMM UI.

View the SMM UI logs on the host.

Limitations

CDPD-36422: 1MB flow.snapshot freezes Safari

While importing large connector configurations, flow.snapshots reduces the usability of the Streams Messaging Manager when using Safari browser.

Use a different browser (Chrome/Firefox/Edge).

Known Issues in Streams Replication Manager

Learn about the known issues in Streams Replication Manager, the impact or changes to the functionality, and the workaround.

Known Issues

CDPD-22089: SRM does not sync re-created source topics until the offsets have caught up with target topic

Messages written to topics that were deleted and re-created are not replicated until the source topic reaches the same offset as the target topic. For example, if at the time of deletion and re-creation there are a 100 messages on the source and target clusters, new messages will only get replicated once the re-created source topic has 100 messages. This leads to messages being lost.

None

CDPD-11079: Blacklisted topics appear in the list of replicated topics

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

will also be visible in the SMM UI. However, its Partitions and Consumer Groups will be 0, its Throughput, Replication Latency and Checkpoint Latency will show N/A.

None

CDPD-30275: SRM may automatically re-create deleted topics on target clusters

If `auto.create.topics.enable` is enabled, deleted topics might get automatically re-created on target clusters. This is a timing issue. It only occurs if remote topics are deleted while the replication of the topic is still ongoing.

1. Remove the topic from the topic allowlist with `srm-control`. For example:

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

2. Wait until SRM is no longer replicating the topic.
3. Delete the remote topic in the target cluster.

OPSAPS-67772: SRM Service metrics processing fails when the noexec option is enabled for /tmp

The SRM Service role uses `/tmp` to extract RocksDB `.so` files, which are required for metrics processing to function. If the `noexec` option is enabled for the `/tmp` directory, the SRM Service role is not able to load the required RocksDB files. This results in metrics processing failing.

1. In Cloudera Manager, select the SRM service and go to Configuration.
2. Add the following to SRM Service Environment Advanced Configuration Snippet (Safety Valve). Do this for all SRM Service role instances.

```
ROCKSDB_SHAREDLIB_DIR=[***PATH***]
```

Replace `[***PATH***]` with a directory that is not `/tmp`.

OPSAPS-67738: SRM Service role's Remote Querying feature does not work when the noexec option is enabled for /tmp

The SRM Service role puts the Netty native libraries into the `/tmp` directory. As a result, if the `noexec` option is enabled for the `/tmp` directory, the Remote Querying feature will fail to function.

1. In Cloudera Manager, select the SRM service and go to Configuration.
2. Add the following to `SRM_JVM_PERF_OPTS`.

```
-Dio.netty.native.workdir=[***PATH***]
```

Replace `[***PATH***]` with a directory that is not `/tmp`.

OPSAPS-67742: The SRM Service role fails to start if properties are added to Additional Configs For Streams Application Running Inside SRM Service

Configuring the SRM Service role's internal Kafka Streams application is not possible. If you add any properties to Using the Additional Configs For Streams Application Running Inside SRM Service, the SRM Service role fails to start. If you are affected by this issue, an exception similar to the following will be present in the SRM Service role's `stderr.log`:

```
o.dropwizard.configuration.ConfigurationParsingException: /var/run/cloudera-scm-agent/process/132-streams_replication_manager-STREAMS_REPLICATION_MANAGER_SERVICE/srm-service.yaml has an error:
* Malformed YAML at line: 66, column: 49; mapping values are not allowed here in 'reader', line 65, column 48:
.
.
```

None

CDPD-60426: Configuration changes are lost following a rolling restart of the service

In certain cases, SRM might fail to apply configuration updates if the service is restarted with a rolling restart. In a case like this, configuration changes are ignored without any warning or indication. This issue also affects rolling upgrades.

When restarting the service, use `Actions Restart` instead of `Actions Rolling Restart` after making configuration changes. When upgrading a cluster, ensure that SRM is not restarted with a rolling restart.

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

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

Known Issues in MapReduce, Apache Hadoop YARN, and YARN Queue Manager

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

Known Issues

CDPD-46685 Nodemanager logs are filled with logs similar to: 2022-11-28 03:42:39,587 WARN org.apache.hadoop.ipc.Client: Address change detected. Old: deh-34631355-niv-master1.e2e-797.dze1-y40r.int.cldr.work/10.114.128.84:8031 New: deh-34631355-niv-master1.e2e-797.dze1-y40r.int.cldr.work/10.114.128.63:8031 2022-11-28 03:43:01,425 WARN org.apache.hadoop.ipc.Client: Address change detected. Old: deh-34631355-niv-master0.e2e-797.dze1-y40r.int.cldr.work/10.114.128.79:8031 New: deh-34631355-niv-master0.e2e-797.dze1-y40r.int.cldr.work/10.114.128.65:8031.

Restart all YARN NodeManagers, they should come up without issues and Cloudera Manager should recognize them as healthy nodes once the status of them is refreshed upon restart.

YARN cannot start if Kerberos principal name is changed

If the Kerberos principal name is changed in Cloudera Manager after launch, YARN will not be able to start. In such case the keytabs can be correctly generated but YARN cannot access ZooKeeper with the new Kerberos principal name and old ACLs.

There are two possible workarounds:

- Delete the znode and restart the YARN service.
- Use the `reset ZK ACLs` command. This also sets the znodes below `/rmstore/ZKRMStateRoot` to `world:anyone:cdrwa` which is less secure.

Third party applications do not launch if MapReduce framework path is not included in the client configuration

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

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

JobHistory URL mismatch after server relocation

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

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

CDH-6808: Routable IP address required by ResourceManager

ResourceManager requires routable host:port addresses for yarn.resourcemanager.scheduler.address, and does not support using the wildcard 0.0.0.0 address.

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

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

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

To view the history for a killed Spark application, see the Spark HistoryServer web UI instead.

COMPX-3329: Autorestart is not enabled for Queue Manager in Data Hub

In a Data Hub cluster, Queue Manager is installed with autorestart disabled. Hence, if Queue Manager goes down, it will not restart automatically.

If Queue Manager goes down in a Data Hub cluster, you must go to the Cloudera Manager Dashboard and restart the Queue Manager service.

COMPX-5817: Queue Manager UI will not be able to present a view of pre-upgrade queue structure. CM Store is not supported and therefore Yarn will not have any of the pre-upgrade queue structure preserved.

When a Data Hub cluster is deleted, all saved configurations are also deleted. All YARN configurations are saved in CM Store and this is yet to be supported in Data Hub and Cloudera Manager. Hence, the YARN queue structure also will be lost when a Data Hub cluster is deleted or upgraded or restored.

Unsupported Features

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

- Application Timeline Server (ATSv2 and ATsv1)
- Container Resizing
- Distributed or Centralized Allocation of Opportunistic Containers
- Distributed Scheduling
- Docker on YARN (DockerContainerExecutor) on Data Hub clusters
- Fair Scheduler
- GPU support for Docker
- Hadoop Pipes
- Native Services
- Pluggable Scheduler Configuration
- Queue Priority Support
- Reservation REST APIs
- Resource Estimator Service
- Resource Profiles
- (non-Zookeeper) ResourceManager State Store
- Rolling Log Aggregation
- Shared Cache
- YARN Federation
- Moving jobs between queues

Known Issues in Apache Zeppelin

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

CDPD-3090: Due to a typo in the configuration, functionality involving notebook repositories does not work

Due to a missing closing brace, access to the notebook repositories API is blocked by default.

1. From the Cloudera Data Platform Management Console, go to Cloudera Manager for the cluster running Zeppelin.
2. On the Zeppelin configuration page (Zeppelin service Configuration), enter shiro urls in the Search field, and then add the missing closing brace to the notebook-repositories URL, as follows:

```
/api/notebook-repositories/** = authc, roles[{{zeppelin_admin_group}}]
```

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

CDPD-2406: Logout button does not work

Clicking the Logout button in the Zeppelin UI logs you out, but then immediately logs you back in using SSO.

Close the browser.

Known Issues in Apache ZooKeeper

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

Zookeeper-client does not use ZooKeeper TLS/SSL automatically

The command-line tool 'zookeeper-client' is installed to all Cloudera Nodes and it can be used to start the default Java command line ZooKeeper client. However even when ZooKeeper TLS/SSL is enabled, the zookeeper-client command connects to localhost:2181, without using TLS/SSL.

Manually configure the 2182 port, when zookeeper-client connects to a ZooKeeper cluster. The following is an example of connecting to a specific three-node ZooKeeper cluster using TLS/SSL:

```
CLIENT_JVMFLAGS="-Dzookeeper.clientCnxnSocket=org.apache.zookeeper.ClientCnxnSocketNetty -Dzookeeper.ssl.keyStore.location=<path to your configured keystore> -Dzookeeper.ssl.keyStore.password=<the password you configured for the keystore> -Dzookeeper.ssl.trustStore.location=<path to your configured truststore> -Dzookeeper.ssl.trustStore.password=<the password you configured for the truststore> -Dzookeeper.client.secure=true" zookeeper-client -server <your.zookeeper.server-1>:2182,<your.zookeeper.server-2>:2182,<your.zookeeper.server-3>:2182
```

Behavioral Changes In Cloudera Runtime 7.2.17

You can review the changes in certain features or functionalities of components that have resulted in a change in behavior from the previously released version to this version of Cloudera Runtime 7.2.17.

Behavioral Changes in Apache Kafka

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

Summary:

The Cluster Health Guarantee During Rolling Restart property is now set to healthy partitions stay healthy. This change is done so that a higher level of cluster health guarantees are provided by default.

Previous behavior:

The default value of the Cluster Health Guarantee During Rolling Restart property was set to none.

New behavior:

The default value of the Cluster Health Guarantee During Rolling Restart property is set to healthy partitions stay healthy.

Behavioral changes in Apache Hive

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

Summary:

Change in default value of the hive.compute.splits.num.threads property

Previous behavior:

The default value for the hive.compute.splits.num.threads property is set to "10"

New behavior:

The default value for the hive.compute.splits.num.threads property is changed to "64" to address a performance issue for Ranger Raz with specific cloud providers. This change affects only Amazon Web Services (AWS) and Google Cloud Platform (GCP) deployments.

Behavioral changes in Apache Ranger

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

Summary:

Ranger REST API response object will not include properties/fields which are NULL or empty/blank.

Ranger Rest API Previous Behavior:

(Before the commit of RANGER-3948)

Ranger REST API response object will include properties/fields which are not NULL and also properties/fields which are empty/blank.

Properties/fields which are having NULL value will be excluded from response object.

Ranger Rest API New Behavior:

(After the commit of RANGER-3948)

Ranger REST API response object will include properties/fields which are not NULL.

Properties/fields which are having NULL or empty/blank values will be excluded from response object.

Behavioral Changes in Apache Solr

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

Summary:

The default value of the `hbaseindexer.httpserver.disabled` environment parameter changed from false to true.

Previous behavior:

You needed to change the value of the `hbaseindexer.httpserver.disabled` environment parameter to true to switch off the REST interface. This was necessary to prevent use of the `--http` argument when using the `hbase-indexer` command line tool. Using the `--http` argument for the `hbase-indexer` command line tool to invoke Lily indexer through REST API allowed adding/listing/removing indexers with any user without the need for authentication.

New behavior:

The HBase Lily indexer REST API is switched off by default.

Deprecation Notices In Cloudera Runtime 7.2.17

Certain features and functionalities have been removed or deprecated in Cloudera Runtime 7.2.17. You must review these items to understand whether you must modify your existing configuration. You can also learn about the features that will be removed or deprecated in the future release to plan for the required changes.

Terminology

Items in this section are designated as follows:

Deprecated

Technology that Cloudera is removing in a future CDP release. Marking an item as deprecated gives you time to plan for removal in a future CDP release.

Moving

Technology that Cloudera is moving from a future CDP release and is making available through an alternative Cloudera offering or subscription. Marking an item as moving gives you time to plan for removal in a future CDP release and plan for the alternative Cloudera offering or subscription for the technology.

Removed

Technology that Cloudera has removed from CDP and is no longer available or supported as of this release. Take note of technology marked as removed since it can potentially affect your upgrade plans.

Removed Components and Product Capabilities

No components are deprecated or removed in this Cloudera Runtime release.

Please contact Cloudera Support or your Cloudera Account Team if you have any questions.

Deprecation Notices for Apache Kafka

Certain features and functionality in Apache Kafka are deprecated or removed in Cloudera Runtime 7.2.17. You must review these changes along with the information about the features in Kafka that will be removed or deprecated in a future release.



Important: The following list of deprecated and removed items is not exhaustive and only contains items that have a direct and immediate effect on Kafka in CDP. For a full list of deprecation and/or removals in the version Apache Kafka shipped with Runtime, review the *Notable Changes* as well as the *Release Notes* on <https://kafka.apache.org/>.

Deprecated

MirrorMaker (MM1)

MirrorMaker is deprecated. Cloudera recommends that you use Streams Replication Manager (SRM) instead.

--zookeeper

The --zookeeper option is only supported for the kafka-configs tool and should be only used when updating SCRAM Credential configurations. The --zookeeper option is either deprecated in or removed from other Kafka command line tools. Cloudera recommends that you use the --bootstrap-server option instead.

Deprecation Notices for Spark 2

Spark 2 is deprecated in Cloudera Runtime 7.2.17. You'll need to migrate your Spark 2 applications to Spark 3.3.2. You must ensure that your jobs are Spark 3.3.2 compliant as Spark 2 will be deprecated in a future release. Please contact Cloudera Support or your Cloudera Account Team if you have any questions

Deprecated

Spark 2

Since Spark 2 is deprecated in 7.2.17, it is highly encouraged that you migrate to Spark 3.3.2 before you upgrade to a later version. See [Updating Spark 2 applications for Spark 3](#) linked below.

Related Information

[Using Spark 2 applications for Spark 3](#)

Fixed Common Vulnerabilities and Exposures 7.2.17

Common Vulnerabilities and Exposures (CVE) that is fixed in this release.

- [CVE-2022-29580](#) - Google Search app
- [CVE-2022-2048](#) - Eclipse Jetty
- [CVE-2022-2047](#) - Eclipse Jetty
- [CVE-2022-34271](#) - Apache Atlas
- [CVE-2022-36364](#) - Apache Calcite Avatica
- [CVE-2020-26939](#) - Bouncy Castle
- [CVE-2021-42550](#) - Logback
- [CVE-2021-29243](#) - Cloudera Manager API
- [CVE-2021-32482](#) - Cloudera Manager API
- [CVE-2022-45868](#) - H2 Database Engine
- [CVE-2021-35517](#) - Commons-compress
- [CVE-2021-36090](#) - Commons-compress
- [CVE-2018-11771](#) - Commons-compress
- [CVE-2012-5783](#) - Commons-httpclient
- [CVE-2021-37533](#) - Commons-net
- [CVE-2022-26612](#) - Apache Hadoop

- [CVE-2021-41561](#) - Apache Parquet
- [CVE-2022-41853](#) - Hsqldb
- [CVE-2018-18928](#) - Icu4j
- [CVE-2020-10531](#) - Icu4j
- [CVE-2020-21913](#) - Icu4j
- [CVE-2022-37865](#) - Apache ivy
- [CVE-2022-37866](#) - Apache ivy
- [CVE-2020-25649](#) - Jackson-databind
- [CVE-2020-28491](#) - Jackson-dataformat
- [CVE-2021-28165](#) - Eclipse Jetty
- [CVE-2021-28169](#) - Eclipse Jetty
- [CVE-2020-27218](#) - Eclipse Jetty
- [CVE-2021-34428](#) - Eclipse Jetty
- [CVE-2022-36033](#) - Jsoup
- [CVE-2022-21724](#) - Postgresql
- [CVE-2022-26520](#) - Postgresql
- [CVE-2022-31197](#) - Postgresql
- [CVE-2022-41946](#) - Postgresql
- [CVE-2021-42575](#) - OWASP Java HTML Sanitizer
- [CVE-2022-36944](#) - Scala
- [CVE-2022-40152](#) - Woodstox
- [CVE-2022-31777](#) - Apache Spark
- [CVE-2022-22971](#) - Spring Framework
- [CVE-2022-22968](#) - Spring Framework
- [CVE-2022-22970](#) - Spring Framework
- [CVE-2022-42252](#) - Apache Tomcat
- [CVE-2022-34305](#) - Apache Tomcat
- [CVE-2021-3642](#) - Wildfly-elytron
- [CVE-2019-10095](#) - Apache Zeppelin
- [CVE-2021-43138](#) - Async
- [CVE-2020-28469](#) - Glob-parent
- [CVE-2020-7598](#) - Minimist
- [CVE-2022-21803](#) - Nconf