

Cloudera Runtime 7.2.18

Release Notes

Date published: 2023-06-20

Date modified: 2024-02-29

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 7.2.18.....	7
Runtime 7.2.18.0-641.....	8
What's New In Cloudera Runtime 7.2.18.....	25
What's New in Apache Atlas.....	26
What's New in Cloud Connectors.....	26
What's New in Cruise Control.....	26
What's New in Apache HBase.....	27
What's New in Apache Hive.....	27
What's New in Hue.....	28
What's new in Apache Iceberg.....	29
What's New in Apache Impala.....	29
What's New in Apache Kafka.....	33
What's New in Apache Knox.....	36
What's New in Apache Kudu.....	36
What's New in Apache Livy.....	37
What's New in Apache Oozie.....	37
What's New in Apache Phoenix.....	37
What's New in Apache Ranger.....	38
What's New in Schema Registry.....	39
What's new in Apache Solr.....	39
What's New in Apache Spark.....	40
What's New in Sqoop.....	41
What's new in Streams Messaging Manager.....	41
What's New in Streams Replication Manager.....	43
What's New in Apache Hadoop YARN and YARN Queue Manager.....	45
What's New in Apache ZooKeeper.....	46
Unaffected Components in this release.....	46
Fixed Issues In Cloudera Runtime 7.2.18.....	46
Fixed Issues in Atlas.....	46
Fixed Issues in Avro.....	48
Fixed Issues in Cloud Connectors.....	48
Fixed issues in Cruise Control.....	49
Fixed Issues in Apache Hadoop.....	49
Fixed Issues in HBase.....	49
Fixed Issues in HDFS.....	50
Fixed Issues in Apache Hive.....	50
Fixed Issues in Hive Warehouse Connector.....	51
Fixed Issues in Hue.....	51
Fixed Issues in Apache Impala.....	53

Fixed Issues in Apache Kafka.....	54
Fixed Issues in Apache Knox.....	54
Fixed Issues in Apache Kudu.....	54
Fixed Issues in Apache Livy.....	56
Fixed Issues in Apache Oozie.....	56
Fixed Issues in Phoenix.....	58
Fixed Issues in Parquet.....	58
Fixed Issues in Apache Ranger.....	59
Fixed Issues in Schema Registry.....	61
Fixed Issues in Apache Solr.....	62
Fixed Issues in Spark.....	62
Fixed Issues in Spark3.....	63
Fixed Issues in Apache Sqoop.....	64
Fixed Issues in Streams Messaging Manager.....	65
Fixed Issues in Streams Replication Manager.....	65
Fixed Issues in Apache Tez.....	66
Fixed Issues in Apache YARN and YARN Queue Manager.....	66
Fixed Issues in Zeppelin.....	67
Fixed Issues in Apache ZooKeeper.....	67

Known Issues In Cloudera Runtime 7.2.18..... 68

Known Issues in Apache Atlas.....	68
Known Issues in Apache Avro.....	74
Known Issues in Cloud Connectors.....	74
Known issues in Cruise Control.....	74
Known Issues in Apache HBase.....	75
Known Issues in HDFS.....	75
Known Issues in Apache Hive.....	76
Known Issues in Hue.....	77
Known Issues Iceberg.....	80
Known Issues in Apache Impala.....	80
Known Issues in Apache Kafka.....	86
Known Issues in Apache Knox.....	89
Known Issues in Apache Kudu.....	90
Known Issues in Apache Oozie.....	90
Known Issues in Apache Phoenix.....	90
Known Issues in Apache Ranger.....	90
Known Issues in Schema Registry.....	91
Known Issues in Apache Solr.....	91
Known Issues in Apache Spark.....	96
Known Issues for Apache Sqoop.....	97
Known issues in Streams Messaging Manager.....	97
Known Issues in Streams Replication Manager.....	98
Known Issues in MapReduce, Apache Hadoop YARN, and YARN Queue Manager.....	99
Known Issues in Apache Zeppelin.....	101
Known Issues in Apache ZooKeeper.....	101

Behavioral Changes In Cloudera Runtime 7.2.18..... 102

Behavioral changes in Apache Hive.....	102
Behavioral Changes in Apache Kafka.....	102
Behavioral changes in Apache Ranger.....	102

Deprecation Notices In Cloudera Runtime 7.2.18..... 104

DAS..... 105

Deprecation Notices for Apache Kafka..... 105

Deprecation Notices for Apache Oozie..... 105

Deprecation Notices for Spark 2..... 106

Deprecation Notices for Zeppelin..... 106

Fixed Common Vulnerabilities and Exposures 7.2.18..... 106

Overview

You can review the Release Notes of Cloudera Runtime 7.2.18 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 7.2.18 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	0.11.1.7.2.18.0-641
Apache Atlas	2.2.0.7.1.8.0-641
Apache Calcite	1.25.0.7.2.18.0-641
Apache Avro	1.8.2.7.2.18.0-641
Apache Flink	1.18.0.1.12.0.0
Apache Hadoop (Includes YARN and HDFS)	3.1.1.7.2.18.0-641
Apache HBase	2.4.17.7.2.18.0-641
Apache Hive	3.1.3000.7.2.18.0-641
Apache Iceberg	1.3.1.7.2.18.0-641
Apache Impala	4.0.0.7.2.18.0-641
Apache Kafka	3.4.1.7.2.18.0-641
Apache Knox	2.0.0.7.2.18.0-641
Apache Kudu	1.17.0.7.2.18.0-641
Apache Livy	0.7.2.7.2.18.0-641
Apache MapReduce	3.1.1.7.2.18.0-641
Apache NiFi Apache NiFi Registry	1.25.0.2.2.8.0
Apache NiFi [Technical Preview] Apache NiFi Registry [Technical Preview]	2.0.0.4.2.0.0
Apache Oozie	5.1.0.7.2.18.0-641
Apache ORC	1.8.3.7.2.18.0-641
Apache Parquet	1.12.3.7.2.18.0-641
Apache Phoenix	5.1.3.7.2.18.0-641
Apache Ranger	2.4.0.7.2.18.0-641
Apache Solr	8.11.2.7.2.18.0-641
Apache Spark	2.4.8.7.2.18.0-641
Apache Spark 3	3.4.1.7.2.18.0-641

Component	Version
Apache Sqoop	1.4.7.7.2.18.0-641
Apache Tez	0.9.1.7.2.18.0-641
Apache Zeppelin	0.8.2.7.2.18.0-641
Apache ZooKeeper	3.8.1.7.2.18.0-641

Other Components

Component	Version
Cruise Control	2.5.116.7.2.18.0-551
Data Analytics Studio	1.4.2.7.2.18.0-641
GCS Connector	2.1.2.7.2.18.0-641
HBase Indexer	1.5.0.7.2.18.0-641
Hive Solr Connector	4.0.0.7.2.18.0-641
Hue	4.5.0.7.2.18.0-641
Search	1.0.0.7.2.18.0-641
Schema Registry	0.10.0.7.2.18.0-641
Spark Solr Connector	3.9.0.7.2.18.0-641
Streams Messaging Manager	2.3.0.7.2.18.0-641
Streams Replication Manager	1.1.0.7.2.18.0-641

Connectors and Encryption Components

Component	Version
HBase connectors	1.0.0.7.2.18.0-641
Hive Meta Store (HMS)	1.0.0.7.2.18.0-641
Hive on Tez	1.0.0.7.2.18.0-641
Hive Warehouse Connector	1.0.0.7.2.18.0-641
Spark Atlas Connector	0.1.0.7.2.18.0-641
Spark Schema Registry	1.1.0.7.2.18.0-641

Using the Cloudera Runtime Maven repository 7.2.18

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 7.2.18.0-641

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	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-aws-s3-bridge	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-azure-adls-bridge	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-classification-updater	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-client-common	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-client-v1	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-client-v2	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-common	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-distro	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-docs	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-graphdb-api	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-graphdb-common	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-graphdb-janus	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-hdfs-bridge	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-index-repair-tool	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-intg	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-janusgraph-hbase2	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-notification	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-plugin-classloader	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-repository	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-server-api	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	atlas-testtools	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	hbase-bridge	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	hbase-bridge-shim	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	hbase-testing-util	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	hdfs-model	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	hive-bridge	2.1.0.7.2.18.0-641

Project	groupId	artifactId	version
Atlas	org.apache.atlas	hive-bridge-shim	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	impala-bridge	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	impala-bridge-shim	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	impala-hook-api	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	kafka-bridge	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	kafka-bridge-shim	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	navigator-to-atlas	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	sample-app	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	sqoop-bridge	2.1.0.7.2.18.0-641
Atlas	org.apache.atlas	sqoop-bridge-shim	2.1.0.7.2.18.0-641
Avro	org.apache.avro	avro	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-compiler	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-ipc	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-mapred	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-maven-plugin	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-protobuf	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-service-archetype	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-thrift	1.8.2.7.2.18.0-641
Avro	org.apache.avro	avro-tools	1.8.2.7.2.18.0-641
Avro	org.apache.avro	trevni-avro	1.8.2.7.2.18.0-641
Avro	org.apache.avro	trevni-core	1.8.2.7.2.18.0-641
Calcite	org.apache.calcite	calcite-babel	1.25.0.7.2.18.0-641
Calcite	org.apache.calcite	calcite-core	1.25.0.7.2.18.0-641
Calcite	org.apache.calcite	calcite-druid	1.25.0.7.2.18.0-641
Calcite	org.apache.calcite	calcite-linq4j	1.25.0.7.2.18.0-641
Calcite	org.apache.calcite	calcite-server	1.25.0.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-aliyun	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-annotations	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-archive-logs	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-archives	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-assemblies	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-auth	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-aws	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-azure	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-azure-datalake	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-benchmark	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-build-tools	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-client	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-client-api	3.1.1.7.2.18.0-641

Project	groupId	artifactId	version
Hadoop	org.apache.hadoop	hadoop-client-integration-tests	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-client-minicluster	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-client-runtime	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-cloud-storage	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-common	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-datajoin	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-distcp	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-extras	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-fs2img	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-gridmix	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-hdfs	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-hdfs-client	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-hdfs-httpfs	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-hdfs-native-client	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-hdfs-nfs	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-hdfs-rbf	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-kafka	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-kms	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-app	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-common	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-core	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-hs	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-hs-plugins	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-jobclient	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-nativetask	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-shuffle	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-uploader	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-mapreduce-examples	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-maven-plugins	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-minicluster	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-minikdc	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-nfs	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-openstack	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-resourceestimator	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-rumen	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-sls	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-streaming	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-tools-dist	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-api	3.1.1.7.2.18.0-641

Project	groupId	artifactId	version
Hadoop	org.apache.hadoop	hadoop-yarn-applications-distributedshell	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-applications-unmanaged-am-launcher	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-client	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-common	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-registry	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-applicationhistoryservice	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-common	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-nodemanager	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-resourcemanager	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-router	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-sharedcachemanager	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-tests	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-timeline-pluginstorage	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-client	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-common	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-server-2	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-tests	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-server-web-proxy	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-services-api	3.1.1.7.2.18.0-641
Hadoop	org.apache.hadoop	hadoop-yarn-services-core	3.1.1.7.2.18.0-641
HBase	org.apache.hbase	hbase-annotations	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-asyncfs	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-checkstyle	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-client	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-client-project	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-common	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-endpoint	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-examples	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-external-blockcache	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-hadoop-compat	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-hadoop2-compat	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-hbtop	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-http	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-it	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-logging	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-mapreduce	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-metrics	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-metrics-api	2.4.17.7.2.18.0-641

Project	groupId	artifactId	version
HBase	org.apache.hbase	hbase-procedure	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-protocol	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-protocol-shaded	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-replication	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-resource-bundle	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-rest	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-rsgroup	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-server	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-shaded-client	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-shaded-client-byo-hadoop	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-shaded-client-project	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-shaded-mapreduce	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-shaded-testing-util	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-shaded-testing-util-tester	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-shell	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-testing-util	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-thrift	2.4.17.7.2.18.0-641
HBase	org.apache.hbase	hbase-zookeeper	2.4.17.7.2.18.0-641
Hive	org.apache.hive	catalogd-unit	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-beeline	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-blobstore	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-classification	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-cli	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-common	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-contrib	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-exec	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-hbase-handler	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-hcatalog-it-unit	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-hplsql	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-iceberg-catalog	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-iceberg-handler	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-iceberg-shading	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-impala	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-custom-serde	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-iceberg	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-impala	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-minikdc	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-qfile	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-qfile-kudu	3.1.3000.7.2.18.0-641

Project	groupId	artifactId	version
Hive	org.apache.hive	hive-it-test-serde	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-unit	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-unit-hadoop2	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-it-util	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-jdbc	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-jdbc-handler	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-jmh	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-kudu-handler	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-llap-client	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-llap-common	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-llap-ext-client	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-llap-server	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-llap-tez	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-metastore	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-parser	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-pre-upgrade	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-serde	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-service	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-service-rpc	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-shims	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-standalone-metastore	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-storage-api	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-streaming	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-testutils	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-udf	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	hive-vector-code-gen	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	kafka-handler	3.1.3000.7.2.18.0-641
Hive	org.apache.hive	patched-iceberg-api	patched-1.3.1.7.2.18.0-641-3.1.3000.
Hive	org.apache.hive	patched-iceberg-core	patched-1.3.1.7.2.18.0-641-3.1.3000.
Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector-spark3_2.12	1.0.0.7.2.18.0-641
Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector_2.11	1.0.0.7.2.18.0-641
Kafka	org.apache.kafka	ci	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-api	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-basic-auth-extension	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-cloudera-authorization-extension	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-cloudera-common	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-cloudera-secret-storage	3.4.1.7.2.18.0-641

Project	groupId	artifactId	version
Kafka	org.apache.kafka	connect-cloudera-security-policies	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-file	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-json	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-mirror	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-mirror-client	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-runtime	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	connect-transforms	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	generator	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-clients	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-cloudera-metrics-reporter_2.12	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-cloudera-metrics-reporter_2.13	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-cloudera-plugins	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-examples	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-group-coordinator	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-log4j-appender	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-metadata	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-raft	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-server-common	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-shell	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-storage	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-storage-api	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-examples	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-scala_2.12	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-scala_2.13	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-test-utils	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0100	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0101	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0102	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0110	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-10	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-11	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-20	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-21	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-22	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-23	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-24	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-25	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-26	3.4.1.7.2.18.0-641

Project	groupId	artifactId	version
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-27	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-28	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-30	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-31	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-32	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-33	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka-tools	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka_2.12	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	kafka_2.13	3.4.1.7.2.18.0-641
Kafka	org.apache.kafka	trogdor	3.4.1.7.2.18.0-641
Knox	org.apache.knox	gateway-adapter	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-admin-ui	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-applications	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-cloud-bindings	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-demo-ldap	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-demo-ldap-launcher	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-discovery-ambari	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-discovery-cm	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-docker	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-i18n	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-i18n-logging-log4j	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-i18n-logging-sl4j	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-openapi-ui	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-performance-test	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-ha	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-identity-assertion-common	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-identity-assertion-concat	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-identity-assertion-hadoop-groups	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-identity-assertion-no-doas	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-identity-assertion-pseudo	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-identity-assertion-regex	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-identity-assertion-switchcase	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-jersey	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-rewrite	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-rewrite-common	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-rewrite-func-hostmap-static	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-rewrite-func-inbound-query-param	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-rewrite-func-service-registry	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-rewrite-step-encrypt-uri	2.0.0.7.2.18.0-641

Project	groupId	artifactId	version
Knox	org.apache.knox	gateway-provider-rewrite-step-secure-query	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-authc-anon	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-authz-acls	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-authz-composite	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-clientcert	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-hadoopauth	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-jwt	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-pac4j	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-preauth	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-shiro	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-provider-security-webappsec	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-release	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-server	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-server-launcher	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-server-xforwarded-filter	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-admin	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-as	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-auth	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-definitions	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-hashicorp-vault	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-hbase	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-health	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-hive	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-idbroker	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-impala	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-jkg	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-knoxsso	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-knoxsout	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-knoxtoken	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-livy	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-metadata	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-nifi	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-nifi-registry	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-remoteconfig	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-rm	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-session	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-storm	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-test	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-tgs	2.0.0.7.2.18.0-641

Project	groupId	artifactId	version
Knox	org.apache.knox	gateway-service-vault	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-service-webhdfs	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-shell	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-shell-launcher	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-shell-release	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-shell-samples	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-spi	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-spi-common	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-test	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-test-idbroker	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-test-release-utils	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-test-utils	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-topology-hadoop-xml	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-topology-simple	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-util-common	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-util-configinjector	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-util-launcher	2.0.0.7.2.18.0-641
Knox	org.apache.knox	gateway-util-urltemplate	2.0.0.7.2.18.0-641
Knox	org.apache.knox	hadoop-examples	2.0.0.7.2.18.0-641
Knox	org.apache.knox	knox-cli-launcher	2.0.0.7.2.18.0-641
Knox	org.apache.knox	knox-homepage-ui	2.0.0.7.2.18.0-641
Knox	org.apache.knox	knox-token-generation-ui	2.0.0.7.2.18.0-641
Knox	org.apache.knox	knox-token-management-ui	2.0.0.7.2.18.0-641
Knox	org.apache.knox	knox-webshell-ui	2.0.0.7.2.18.0-641
Knox	org.apache.knox	webhdfs-kerb-test	2.0.0.7.2.18.0-641
Knox	org.apache.knox	webhdfs-test	2.0.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-backup-tools	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-backup2_2.11	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-backup3_2.12	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-client	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-hive	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-spark2-tools_2.11	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-spark2_2.11	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-spark3-tools_2.12	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-spark3_2.12	1.17.0.7.2.18.0-641
Kudu	org.apache.kudu	kudu-test-utils	1.17.0.7.2.18.0-641
Livy	org.apache.livy	livy-api	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-client-common	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-client-http	0.7.23000.7.2.18.0-641

Project	groupId	artifactId	version
Livy	org.apache.livy	livy-core_2.11	0.7.2.7.2.18.0-641
Livy	org.apache.livy	livy-core_2.12	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-examples	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-integration-test	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-repl_2.11	0.7.2.7.2.18.0-641
Livy	org.apache.livy	livy-repl_2.12	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-rsc	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-scala-api_2.11	0.7.2.7.2.18.0-641
Livy	org.apache.livy	livy-scala-api_2.12	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-server	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-test-lib	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-thriftserver	0.7.23000.7.2.18.0-641
Livy	org.apache.livy	livy-thriftserver-session	0.7.23000.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-common	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-icu	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-kuromoji	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-morfologik	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-nori	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-openslp	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-phonetic	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-smartcn	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-analyzers-stempel	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-backward-codecs	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-benchmark	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-classification	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-codecs	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-core	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-demo	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-expressions	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-facet	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-grouping	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-highlighter	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-join	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-memory	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-misc	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-monitor	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-queries	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-queryparser	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-replicator	8.11.2.7.2.18.0-641

Project	groupId	artifactId	version
Lucene	org.apache.lucene	lucene-sandbox	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-spatial-extras	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-spatial3d	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-suggest	8.11.2.7.2.18.0-641
Lucene	org.apache.lucene	lucene-test-framework	8.11.2.7.2.18.0-641
Oozie	org.apache.oozie	oozie-client	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-core	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-distro	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-examples	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-fluent-job-api	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-fluent-job-client	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-server	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-distcp	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-git	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-hcatalog	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-hive	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-hive2	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-oozie	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-spark	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-spark3	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-sqoop	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-sharelib-streaming	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-tools	5.1.0.7.2.18.0-641
Oozie	org.apache.oozie	oozie-zookeeper-security-tests	5.1.0.7.2.18.0-641
ORC	org.apache.orc	orc-core	1.8.3.7.2.18.0-641
ORC	org.apache.orc	orc-examples	1.8.3.7.2.18.0-641
ORC	org.apache.orc	orc-mapreduce	1.8.3.7.2.18.0-641
ORC	org.apache.orc	orc-shims	1.8.3.7.2.18.0-641
ORC	org.apache.orc	orc-tools	1.8.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-avro	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-cascading	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-cascading3	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-column	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-common	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-encoding	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-format-structures	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-generator	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-hadoop	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-hadoop-bundle	1.12.3.7.2.18.0-641

Project	groupId	artifactId	version
Parquet	org.apache.parquet	parquet-jackson	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-pig	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-pig-bundle	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-protobuf	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-scala_2.12	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-thrift	1.12.3.7.2.18.0-641
Parquet	org.apache.parquet	parquet-tools	1.12.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-client-embedded-hbase-2.4	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-client-hbase-2.4	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-core	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.1.6	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.2.5	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.3.0	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.4.0	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.4.1	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.5.0	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.5.4	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-pherf	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-queryserver	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-queryserver-client	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-queryserver-it	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-queryserver-load-balancer	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-queryserver-orchestrator	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-server-hbase-2.4	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix-tracing-webapp	5.1.3.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-flume	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-hive	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-hive-shaded	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-kafka	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-spark	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-spark-shaded	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-spark3	6.0.0.7.2.18.0-641
Phoenix	org.apache.phoenix	phoenix5-spark3-shaded	6.0.0.7.2.18.0-641
Ranger	org.apache.ranger	conditions-enrichers	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	credentialbuilder	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	embeddedwebserver	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	jisql	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ldapconfigcheck	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-adls-plugin	2.4.0.7.2.18.0-641

Project	groupId	artifactId	version
Ranger	org.apache.ranger	ranger-atlas-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-atlas-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-authn	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-common-ha	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-distro	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-examples-distro	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-gs-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-hbase-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-hbase-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-hdfs-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-hdfs-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-hive-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-hive-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-intg	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kafka-connect-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kafka-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kafka-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kms	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kms-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kms-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-knox-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-knox-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kudu-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kylin-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-kylin-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-metrics	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-nifi-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-nifi-registry-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-ozone-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-ozone-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-plugin-classloader	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-plugins-audit	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-plugins-common	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-plugins-cred	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-plugins-installer	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-policymigration	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-raz-adls	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-raz-chained-plugins	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-raz-hook-abfs	2.4.0.7.2.18.0-641

Project	groupId	artifactId	version
Ranger	org.apache.ranger	ranger-raz-hook-s3	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-raz-intg	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-raz-processor	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-raz-s3	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-raz-s3-lib	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-rms-common	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-rms-hive	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-rms-plugins-common	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-rms-webapp	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-s3-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-sampleapp-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-schema-registry-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-solr-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-solr-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-sqoop-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-sqoop-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-storm-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-storm-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-tagsync	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-tools	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-util	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-yarn-plugin	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ranger-yarn-plugin-shim	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	sample-client	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	sampleapp	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	shaded-raz-hook-abfs	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	shaded-raz-hook-s3	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	ugsync-util	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	unixauthclient	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	unixauthservice	2.4.0.7.2.18.0-641
Ranger	org.apache.ranger	unixusersync	2.4.0.7.2.18.0-641
Solr	org.apache.solr	solr-analysis-extras	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-analytics	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-cell	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-core	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-dataimporthandler	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-dataimporthandler-extras	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-gcs-repository	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-jaegertracer-configurator	8.11.2.7.2.18.0-641

Project	groupId	artifactId	version
Solr	org.apache.solr	solr-langid	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-ltr	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-prometheus-exporter	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-s3-repository	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-security-util	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-solrj	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-test-framework	8.11.2.7.2.18.0-641
Solr	org.apache.solr	solr-velocity	8.11.2.7.2.18.0-641
Spark	org.apache.spark	spark-avro_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-avro_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-catalyst_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-catalyst_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-connect-client-jvm_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-connect-common_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-connect_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-core_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-core_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-graphx_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-graphx_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-hadoop-cloud_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-hadoop-cloud_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-hive_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-hive_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-kubernetes_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-kubernetes_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-kvstore_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-kvstore_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-launcher_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-launcher_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-mllib-local_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-mllib-local_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-mllib_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-mllib_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-network-common_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-network-common_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-network-shuffle_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-network-shuffle_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-network-yarn_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-network-yarn_2.12	3.4.1.7.2.18.0-641

Project	groupId	artifactId	version
Spark	org.apache.spark	spark-protobuf_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-repl_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-repl_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-shaded-raz	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-sketch_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-sketch_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-sql-kafka-0-10_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-sql-kafka-0-10_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-sql_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-sql_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-streaming-kafka-0-10_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-streaming-kafka-0-10_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-streaming_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-streaming_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-tags_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-tags_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-token-provider-kafka-0-10_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-token-provider-kafka-0-10_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-unsafe_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-unsafe_2.12	3.4.1.7.2.18.0-641
Spark	org.apache.spark	spark-yarn_2.11	2.4.8.7.2.18.0-641
Spark	org.apache.spark	spark-yarn_2.12	3.4.1.7.2.18.0-641
Sqoop	org.apache.sqoop	sqoop	1.4.7.7.2.18.0-641
Sqoop	org.apache.sqoop	sqoop-test	1.4.7.7.2.18.0-641
Tez	org.apache.tez	hadoop-shim	0.9.1.7.2.18.0-641
Tez	org.apache.tez	hadoop-shim-2.8	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-api	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-aux-services	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-common	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-dag	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-examples	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-ext-service-tests	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-history-parser	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-javadoc-tools	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-job-analyzer	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-mapreduce	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-protobuf-history-plugin	0.9.1.7.2.18.0-641

Project	groupId	artifactId	version
Tez	org.apache.tez	tez-runtime-internals	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-runtime-library	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-tests	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-yarn-timeline-cache-plugin	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-yarn-timeline-history	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-yarn-timeline-history-with-acls	0.9.1.7.2.18.0-641
Tez	org.apache.tez	tez-yarn-timeline-history-with-fs	0.9.1.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-angular	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-display	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-interpreter	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-jdbc	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-jupyter	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-livy	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-markdown	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-server	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-shaded-raz	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-shell	0.8.2.7.2.18.0-641
Zeppelin	org.apache.zeppelin	zeppelin-zengine	0.8.2.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-fatjar	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-loggraph	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-rest	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-it	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-jute	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-prometheus-metrics	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-election	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-lock	3.8.1.7.2.18.0-641
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-queue	3.8.1.7.2.18.0-641

What's New In Cloudera Runtime 7.2.18

You must be aware of the additional functionalities and improvements to features of components in Cloudera Runtime 7.2.18. Learn how the new features and improvements benefit you.

Zero Downtime OS upgrade support for service specific safe stop command

Cloudera Manager introduces Zero Downtime OS upgrade feature to improve the upgrade process.

The CDP Runtime services can now intelligently postpone the stopping of role instances until the appropriate time. Thus, ensuring continuous service availability throughout the upgrade process. This improvised process allows for a seamless and uninterrupted availability of services during the OS upgrades.



Important: Currently only KAFKA service supports service availability during OS upgrades.

What's New in Apache Atlas

Learn about the new features of Apache Atlas in Cloudera Runtime 7.2.18.

Iceberg support for Atlas

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

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

Storage reduction for Atlas

Audit aging reduces the existing audit data in the Atlas system which is based on the end user criteria and configuration changes that users can manage.

See [Storage reduction for Atlas](#) for more information.

What's New in Cloud Connectors

Learn about the new features of Cloud Connectors in Cloudera Runtime 7.2.18.

Support for Amazon S3 Express One Zone Storage

Support for Amazon S3 Express One Zone is added. The Amazon S3 Express One Zone is a single-Availability Zone and high-performance storage class that delivers consistent single-digit millisecond data access. A specific Availability Zone can be selected within an AWS Region to store your data. This enables you to have your storage and compute resources in the same Availability Zone to optimize performance, lower compute costs and run workloads faster.

Vectored IO support

Support for Hadoop Vectored IO API is added for ORC and Parquet file formats. The S3A connector offers a customized implementation that enables parallel and asynchronous reading of different data blocks.

Migration to AWS V2 SDK


AWS V2 Java SDK is used for communicating with AWS services, which includes storage through the S3A connector.

What's New in Cruise Control

Learn about the new features of Cruise Control in Cloudera Runtime 7.2.18.

Cruise Control is added to Streams Messaging Manager UI

A new page is added to Streams Messaging Manager to monitor the Kafka cluster state and rebalancing process with Cruise Control. The Cruise Control User Interface (UI) enables you to review and configure the rebalancing of Kafka clusters through dashboards and a rebalancing wizard. The available goals and anomaly detectors are based on the

Cloudera Manager configurations of Cruise Control. You can access Cruise Control from SMM using the  on the navigation sidebar.

For more information about Cruise Control in SMM, see [Monitoring and managing Kafka cluster rebalancing](#).

What's New in Apache HBase

Learn about the new features of HBase in Cloudera Runtime 7.2.18.

HBase supports load balancing using a cache-aware load balancer

The HBase balancer now supports the cache-aware load balancer that enhances the capability of HBase to enable the balancer to consider the cache allocation of each region on region servers while calculating a new assignment plan. This balancer also uses the region or region server cache allocation information reported by the region servers to calculate the percentage of HFiles cached for each region on the hosting server, and then use that as an additional factor while deciding an optimal new assignment plan.

HBase supports Snappy with /tmp directory mounted with noexec option

In Cloudera Manager, the Snappy temporary directory configuration item is added to HBase Master and HBase RegionServer to allow Snappy compression when /tmp directory is mounted with noexec option.

HBase supports Netty native libraries with /tmp directory mounted with noexec option

In Cloudera Manager, the Netty native library working directory configuration item is added to HBase Master and HBase RegionServer to support HBase with /tmp directory mounted with noexec option.

HBase shows cached percentage for region data on RegionServer UI

An important feature for Cloudera Operational Database (COD) over S3 with ephemeral cache is the process of warming up the cache at region opening (also known as *cache prefetch*). The goal is to load the most of the dataset before any client reads, so that a reduced latency and optimal performance can be achieved for the application requests. This *prefetch* process takes several hours on very large datasets, and the operators might want to monitor the progress of this cache loading. To handle this, HBase has introduced new metrics about the percentage of individual regions data currently cached, and it also added this information to the Storefile Metrics tab in the Regions section of the RegionServer UI.

Related Apache JIRA: [HBASE-28246](#)

Region Name	Num. Stores	Num. Storefiles	Storefile Size Uncompressed	Storefile Size	Index Size	Bloom Size	Data Locality	% Cached
hbase:namespace,1710343972022.123ea5d1573867d44a01e1460a5dda68.	1	1	1 MB	1 MB	1 KB	1 KB	1.0	100.00%
my-user-table.A1009037-128-dup1709753704318,1709757279182.64a3c9c7f85002d1c88baff16c5e7527.	2	2	201 MB	32 MB	31 KB	256 KB	1.0	100.00%
my-user-table.A1009093-dup1709748054721,1709759852911.a34c57f4fe6ad1f778e0f4e7c343579d.	2	2	405 MB	66 MB	63 KB	512 KB	1.0	100.00%
my-user-table.A1009238-dup1709753529701,1709758902574.cd99cd790caa93522d9beb313d14b0c0.	2	2	402 MB	65 MB	63 KB	512 KB	1.0	100.00%
my-user-table.A1009226-208-dup1709749478903,1709756783373.9e1abc072025cabfbfc7c945212525a1.	2	2	399 MB	65 MB	62 KB	512 KB	1.0	98.46%

HBase supports disabling the caching for the individual column families

In some use cases, not all tables in the dataset have the same SLA requirements. If the total cache capacity is much smaller than the whole dataset, an alternative is to restrict the cache usage by the tables with critical response times. In HBase, you can now implement this by disabling the cache on individual column families.

On an hbase shell, perform the following `alter` command for each column family that does not require caching.

```
alter 'NAMESPACE:TABLENAME', {NAME=>'CF_NAME', BLOCKCACHE => 'false'}
```

What's New in Apache Hive

Learn about the new features of Hive in Cloudera Runtime 7.2.18.

HiveServer graceful shutdown

You can now configure the graceful shutdown timeout property for HiveServer (HS2), which ensures that HS2 waits for a specified time period before shutting down, thereby allowing queries that are already running to complete before HS2 stops. For more information, see [Configuring graceful shutdown property for HiveServer](#).

Support for column histogram statistics

In this release, when you generate column statistics in Hive, you can create histogram statistics on columns. By default, the histogram stats are not created. You enable generation of histogram statistics by setting a Hive property: `set hive.stats.kll.enable = true;`

You can then run the ANALYZE command as usual:

```
ANALYZE TABLE [table_name] COMPUTE STATISTICS for COLUMNS [comma_separated_column_list];
```

Histogram statistics are supported for numeric data types, date, timestamp and boolean types but not for string/varchar/char columns. Histograms are used to estimate selectivity of range predicates (predicates involving <, <=, >, >= and BETWEEN). The better selectivity estimate allows the optimizer to generate more optimal query plans and improve performance for such queries.

What's New in Hue

Learn about the new features of Hue in Cloudera Runtime 7.2.18.

Hue supports natural language query processing (Preview)

Hue leverages the power of Large Language Models (LLM) to help you generate SQL queries from natural language prompts and also provides options to optimize, explain, and fix queries, ensuring efficiency and accuracy in data retrieval and manipulation. You can use several AI services and models such as OpenAI's GPT service, Amazon Bedrock, and Azure's OpenAI service to run the Hue SQL AI assistant. See [SQL AI Assistant in Cloudera DataHub](#).

Ability to access GS buckets from Hue with RAZ

For better security and ease of use for users, you can create per-user home directories within your Google Cloud Storage bucket (GS) and grant fine-grained access to these user directories from the GS File Browser in Hue. To enable fine-grained access from the Hue GS File Browser, you must enable Ranger Authorization Service (RAZ) when registering your GCP environment with CDP. For more information, see [Using Google Cloud Storage with Hue in Cloudera DataHub](#).

Hue supports Hive Hybrid Procedural SQL

You can run Hive Hybrid Procedural SQL (HPL/SQL) using the Hue query editor. To enable the HPL/SQL interpreter, see [Enabling stored procedures for Hive in DataHub](#). To run stored procedures from Hue, see [How to run a stored procedure from Hue](#).

Ability to control caching behavior of the web page

You can enable web page caching to ensure that your browser fetches the latest resources while you are exploring data using Hue. Hue uses Cache-Control, Pragma, and Expires HTTP headers. To enable cache control, see [Enabling cache-control HTTP headers when using Hue](#).

Ability to create tables by importing files using Hue Importer

You can create Hive, Impala, and Iceberg tables by importing CSV or XLSX files in Hue using Hue Importer. Using Importer you can either upload files up to 200 KB from your local computer and then create tables or browse S3,

ADLS Gen2 storage, or Google Cloud Storage buckets and import the files into Hue. For more information, see [Creating tables in Hue by importing files](#).

Improvements

Hue with RAZ supports High Availability on AWS, Azure, and Google Cloud

You can now specify a comma-separated list of URLs in the `api_url` property while enabling S3, ABFS, or GS File Browser for Hue with RAZ in DataHub. For example:

```
api_url=https://[***INSTANCE-1***]:6082/,https://[***INSTANCE-2***]:6082/
```

Ability to open Oozie tasks and workflows in a new tab using the middle mouse button (scroll wheel)

The functionality to open Oozie-related links in a new tab by clicking the link with the middle button has been restored. You can also press Ctrl (Windows) or Command (MacOS) and click the link to open Oozie tasks in a new tab.

Security improvements

- Cloudera has fixed several CVEs and cross-site scripting vulnerabilities in Hue.
- Knox and Hue Load Balancer hostnames are automatically added to `trusted_origin` and `knox_proxyhost` configuration properties.
- SAML certificates can be created using passphrases.
- SAML sign-in algorithm has been updated from SHA1 to SHA256.
- The cryptography Python module has been upgraded from `cryptography==36.0.1` to `cryptography==41.0.1`.

What's new in Apache Iceberg

Learn about the new features of Iceberg in Cloudera Runtime 7.2.18.

General availability of Iceberg and Atlas integration

[Iceberg in Apache Atlas](#) helps you identify the Iceberg tables to scan data and provide lineage support. Learn how Atlas works with Iceberg and what schema evolution, partition specification, partition evolution are with examples.

Hive to Iceberg table migration from Impala

In this release, you can use Impala, as well as Hive, to migrate a Hive table to Iceberg tables. You use the ALTER TABLE statement. Syntax is described in [Migrate Hive table to Iceberg feature](#) and a step-by-step procedure is covered in [Migrating a Hive table to Iceberg](#).

Iceberg position delete feature support

In this release, Impala, in addition to Hive, can [delete Iceberg V2 tables](#) using position delete files, a format defined by the Iceberg Spec. A position delete query evaluates rows from one table against a WHERE clause, and delete all the rows that match WHERE conditions.

What's New in Apache Impala

Learn about the new features of Impala in Cloudera Runtime 7.2.18.

Removing self-generated events

In previous releases, metadata consistency issues resulted in query failures. This occurred because the metadata updates from various coordinators couldn't distinguish between events generated by the coordinator itself and those

generated by a different coordinator. This release addresses this issue by introducing a coordinator flag to each event. When processing these events, we now examine the coordinator flag to determine whether to ignore the event or proceed accordingly, resolving the inconsistency and preventing query failures.

Impala WebUI improvements

This release introduces significant enhancements to the Impala daemon's Web UI, providing users with additional insights into the system's performance:

Backends Start Time and Version:

- In large clusters, the Impala daemon's Web UI now allows you to easily access and view the start time and version details for all backends.

Query Performance Characteristics:

- Gain deeper insights into query execution with a detailed report on how a query was executed. The built-in web server's UI features a [Gantt chart](#) timeline, serving as an alternative to the PROFILE command. This graphical display in the Web UI renders timing information and dependencies.

Export Query Plan and Timeline:

- As an alternative to the PROFILE command's profile download page, this release introduces support for exporting the graphical query plan and downloading the timeline in SVG/HTML format. Exporting these elements clears memory resources consumed from the ObjectURLs.

Historical/In-flight Query Performance:

- The query list and query details page now offer the capability to analyze historical or in-flight query performance. Users can access information such as memory consumption, data read, and other relevant details about each query.

JWT auth for Impala

Authentication is a crucial mechanism to secure connections to Impala, ensuring that only designated hosts and users can access the system. To implement JSON Web Token (JWT) authentication for Impala, follow these steps:

Configuration in CDP with Cloudera Manager:

- Begin by configuring JWT authentication in [Cloudera Data Platform \(CDP\) using Cloudera Manager](#). This involves setting up the necessary parameters and security settings to enable JWT authentication.

Client Authentication:

- Once JWT authentication is configured, clients—such as the Impala shell—can authenticate to Impala using a JWT instead of the traditional username/password combination. This enhances security and provides an alternative, token-based approach to authentication.

By adopting JWT authentication, Impala ensures a more secure and efficient authentication process for connecting hosts and users. This method offers a modern and flexible alternative to the conventional username/password authentication mechanism.

TPC-DS performance improvements

This release incorporates several enhancements across the planner and executor components to elevate query performance and align with the TPC decision support (TPC-DS) benchmark standards. The key improvements include:

Cardinality Estimation for Joins:

- Significantly enhances cardinality estimation for [joins involving multiple conjuncts](#), leading to more accurate query execution plans and improved performance.

Memory Estimation for Aggregation Nodes:

- Introduces new query options specifically designed to enhance memory [estimation for aggregation nodes](#). This optimization contributes to more efficient memory utilization during query execution.

Planner changes for CPU usage:

- Implements changes in the query planner to enhance parallel sizing and resource estimation, catering to workload-aware autoscaling. The introduced query options allow users to fine-tune these settings for improved CPU utilization and overall performance. This feature enables the global activation of multi-threaded queries, offering enhanced scalability.

Late Materialization of Columns:

- [Introduces late materialization](#), a feature optimizing certain queries on Parquet tables. This optimization minimizes table scanning by materializing only the relevant data, thereby improving query response times.

These improvements collectively contribute to a more robust and efficient Impala system, ensuring optimal performance and compliance with TPC-DS benchmark standards. Users can leverage the new query options for tuning purposes and take advantage of late materialization to enhance the processing of queries on Parquet tables.

Resetting all query options

The [unset all](#) command provides a convenient way to reset all query options. This functionality becomes particularly valuable in scenarios where connections are reused, such as when utilizing a connection pool. By executing UNSET ALL, all query options are unset, allowing for a clean slate and ensuring that subsequent queries operate with default settings. This capability enhances flexibility and efficiency, especially in connection pool scenarios where a fresh start for query options is desired.

Limited support for Hive Generic UDFs

In this release, support for the second generation of Hive User-Defined Functions (UDFs), known as GenericUDFs, is introduced. However, it comes with certain [limitations](#) that users should be aware of:

Decimal Types Not Supported:

- GenericUDFs in this release do not provide support for decimal types, and their usage with such data types may lead to limitations or errors.

Complex Types Not Supported:

- The support for GenericUDFs is limited, and complex types are not currently supported. Users should be mindful of this restriction when working with UDFs that involve complex data structures.

Functions Not Extracted from JAR Files:

- Unlike other UDF types, GenericUDFs do not automatically extract functions from JAR files. Users need to manually manage and ensure that the required functions are appropriately included for use.

Non-Permanent Nature:

- GenericUDFs created in this release are not permanent and will not persist across server restarts. Recreating them is necessary after each server restart to maintain functionality.

These limitations highlight considerations for users employing GenericUDFs in their workflows. It is advised to evaluate these constraints and plan accordingly when incorporating GenericUDFs into Hive queries.

Printing Query Results in Vertical Format

In the latest update, Impala-shell introduces a new command option '-E' or '--vertical' to facilitate the printing of [query results in a vertical format](#). This provides users with a more streamlined and readable display of query outputs.

Retrieving the Data File Name

Impala now offers support for including [a virtual column in a standard SELECT statement](#). By using the following syntax: `SELECT INPUT__FILE__NAME FROM <tablename>`, users can effortlessly retrieve the name of the data

file associated with the actual row stored in a table. This enhancement provides valuable insights into the underlying data organization.

Resolving ORC Columns by Names

In previous releases, Impala resolved ORC columns based on index. With the introduction of this release, [a new query option](#), `ORC_SCHEMA_RESOLUTION`, is now available. This option allows users to resolve ORC columns by names, offering a more flexible and intuitive approach to working with ORC data.

Reading and Writing Parquet Bloom Filters

Introducing a performance optimization feature in Impala — the [Parquet bloom filter](#). This feature enables rapid and memory-efficient determination of whether the desired data is present in a file. Users can now benefit from enhanced efficiency when working with Parquet files.

BYTES Function Support

Impala now incorporates support for the [BYTES\(\) function](#). This function efficiently returns the number of bytes contained within a byte string. Users can leverage this functionality to gain insights into the size of byte strings within their data.

Min/Max Filtering in Impala

With the utilization of the Parquet format, Impala introduces the capability to perform min/max filtering. Users can now execute queries to identify the [minimum or maximum](#) value for a column within various levels such as partition, row group, page, or row. This enhancement provides a more granular and targeted approach to data analysis.

DDL Support for Bucketed Tables

In the latest release, Impala introduces [Data Definition Language \(DDL\) support for bucketed tables](#). This feature enables users to optimize query performance by creating tables with bucketing. Leveraging the `CLUSTER BY` clause, this functionality facilitates the partitioning of data into smaller, more manageable segments based on specified columns. This enhancement contributes to improved query efficiency and data organization.

Support for Collections of Fixed-Length Types as Non-Passthrough Children of Unions

In this release, Impala introduces support for collections of fixed-length types as non-passthrough children of unions. While plain `UNION`s are not yet supported for any collections, `UNION ALL` operations are fully supported. Users can take advantage of this feature to combine and analyze data efficiently within complex queries.

Example:

```
select id, int_array from complextypestbl
union all select cast(id as tinyint), int_array from complextypestbl
```

Support for ORDER BY in Collections of Fixed-Length Types in SELECT List

With this release, Impala now supports collections of [fixed-length types](#) in the sorting tuple. Although sorting directly by these collection columns is not permitted, they can be included in the `SELECT` list alongside other columns by which sorting is applied. This enhancement provides users with greater flexibility in organizing and presenting query results.

Support for Complex Types in SELECT List

In this release, Impala introduces comprehensive support for complex types in the `SELECT` list. While collections and structs were previously supported, the nesting and mixing of complex types were not. Now, users can leverage the flexibility of embedding complex types into other complex types, providing enhanced versatility in query results. For detailed information and any limitations, refer to the "Allowing Embedding Complex Types into Other Complex Types" section in the [Complex types](#) documentation.

Structs in SELECT List with Beeswax

In previous releases, structs in the select list were limited to the HS2 protocol. With this release, the support for structs in the select list is extended to Beeswax as well. Users can now benefit from using structs in the select list when interacting with Beeswax, improving the consistency of functionality across different protocols.

Query Hints for Table Cardinalities

Impala now offers improved control over query planning with the introduction of query hints for table cardinalities. Previously, Impala relied on simple estimation to compute selectivity, which could deviate significantly from actual values for certain predicates, leading to suboptimal query plans. With the addition of [a new query hint](#), 'SELECTIVITY', users can now specify selectivity values for predicates, enabling more accurate query planning and better overall performance.

What's New in Apache Kafka

Learn about the new features of Apache Kafka in Cloudera Runtime 7.2.18.

Rebase on Kafka 3.4.1

Kafka shipped with this version of Cloudera Runtime is based on Apache Kafka 3.4.1. 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](#)
- [3.4.1](#)

Kafka log directory monitoring improvements

A new Cloudera Manager chart, trigger, and action is added for the Kafka service. These assist you in monitoring the log directory space of the Kafka Brokers, and enable you to prevent Kafka disks from filling up.

The chart is called Log Directory Free Capacity. It shows the capacity of each Kafka Broker log directory.

The trigger is called Broker Log Directory Free Capacity Check. It is triggered if the capacity of any log directory falls below 10%. The trigger is automatically created for all newly deployed Kafka services, but must be created with the Create Kafka Log Directory Free Capacity Check action for existing services following an upgrade.

The chart and trigger are available on the [Kafka service Status](#) page. The action is available in [Kafka service Actions](#).

Kafka is safely stopped during operating system upgrades

During OS upgrades, Cloudera Manager now ensures that Kafka brokers are safely stopped. Specifically, Cloudera Manager now performs a rolling restart check before stopping a broker. This ensures that the Kafka service stays healthy during the upgrade. The level of health guarantee that Cloudera Manager ensures is determined by the restart check type set in the Cluster Health Guarantee During Rolling Restart Kafka property. Cloudera recommends that you set this property to all partitions stay healthy to avoid service outages. For more information, see [Rolling restart checks](#).

useSubjectCredsOnly set to true by default in Kafka Connect

In previous versions, the `javax.security.auth.useSubjectCredsOnly` JVM property was set to false in Kafka Connect. Because of this, connectors running with an invalid or no JAAS configuration could use the credentials of other connectors to establish connections. Starting with this release, `useSubjectCredsOnly` is set to true by default. As a result, connectors are required to use their own credentials.

This default change is true for newly provisioned clusters. On upgraded clusters, `useSubjectCredsOnly` remains set to false to ensure backwards compatibility. If you are migrating connectors from a cluster running a previous version of Runtime to a new cluster running 7.2.18 or later, you must ensure that credentials are added to the connector configuration when migrated. Otherwise, migrated connectors may not work on the new cluster.

In addition to the default value change, a new Kafka Connect property is introduced in Cloudera Manager that you can use to set `useSubjectCredsOnly`. The property is called `Add Use Subject Credentials Only JVM Option With True Value`. Setting this property to false does not expressly set `useSubjectCredsOnly` to false. Instead, it sets `useSubjectCredsOnly` to the cluster default value.

Kafka Connect metrics reporter security configurable in Cloudera Manager

New, dedicated Cloudera Manager properties are introduced for the security configuration of the Kafka Connect metrics reporter. As a result, you are no longer required to use advanced security snippets if you want to secure the metrics reporter and its endpoint. The new properties introduced are as follows:

- Secure Jetty Metrics Port
- Enable Basic Authentication for Metrics Reporter
- Jetty Metrics User Name
- Jetty Metrics Password

A dedicated property to enable TLS/SSL for the metrics reporter is not available. Instead, you must select `Enable TLS/SSL for Kafka Connect` which enables TLS/SSL for the Kafka Connect role including the metrics reporter. For more information regarding these properties, see [Cloudera Manager Configuration Properties Reference](#).

As a result of these changes, the setup steps required to configure Prometheus as the metrics store for SMM are changed. For updated deployment instructions, see [Setting up Prometheus for Streams Messaging Manager](#).

Kafka load balancer is automatically configured with the LDAP handler if LDAP authentication is configured

When a load balancer and LDAP authentication is configured for Kafka, the PLAIN mechanism is automatically added to the enabled authentication mechanisms of the load balancer listener. Additionally, the load balancer is automatically configured to use `LdapPlainServerCallbackHandler` as the callback handler.

Kafka Connect now supports Kerberos auth-to-local (ATL) rules with SPNEGO authentication

Kafka Connect now uses the cluster-wide Kerberos auth-to-local (ATL) rules by default. A new configuration property called `Kafka Connect SPNEGO Auth To Local Rules` is introduced. This property is used to manually specify the ATL rules. During an upgrade, the property is set to `DEFAULT` to ensure backward compatibility. Following an upgrade, if you want to use the cluster-wide rules, clear the existing value from the `Kafka Connect SPNEGO Auth To Local Rules` property.

Debezium connector version update

All Debezium connectors shipped with Cloudera Runtime are upgraded to version 1.9.7. Existing instances of the connectors are automatically upgraded to the new version during cluster upgrade. Deploying the previously shipped version of the connector is not possible. For more information see [Kafka Connectors in Runtime](#) or the [Debezium documentation](#).

Persistent MQTT sessions support for the MQTT Source connector

Version 1.1.0 of the MQTT Source connector is released. The connector now supports MQTT persistent sessions. This enables the connector to resume (persist) a previous session with an MQTT broker after a session is interrupted. Enabling this feature can ensure that no messages are lost if the connector is momentarily stopped or if the network connection is interrupted.

To support persistent sessions, the following new properties are introduced:

- MQTT Client ID

This property specifies the MQTT client ID that the connector uses.

- MQTT Clean Session

This property controls whether the connector should start clean or persistent sessions. Set this property to false to enable persistent sessions.

Existing connectors will continue to function, upgrading them, however, is not possible. If you want to use the new version of the connector, you must deploy a new instance of the connector. For more information, see [MQTT Source connector](#) and [MQTT Source properties reference](#).

Parquet support for the S3 Sink connector

Version 2.0.0 of the S3 Sink connector is released. The connector now supports Parquet as an output file data format. The following property changes are made to support Parquet:

- A new property, Parquet Compression Type, is added.

This property specifies the compression type used for writing Parquet files. Accepted values are UNCOMPRESSED, SNAPPY, GZIP, LZO, BROTLI, LZ4, and ZSTD.

- The Output File Data Format property now accepts Parquet as a value.

Existing connectors will continue to function, upgrading them, however, is not possible. If you want to use the new version of the connector, you must deploy a new instance of the connector.

For more information, see [S3 Sink connector](#) and [S3 Sink properties reference](#).

Support schema ID encoding in the payload or message header in Stateless NiFi connectors

The Kafka Connect connectors powered by Stateless NiFi that support record processing are updated to support content-encoded schema references for Avro messages. These connectors now properly support integration with Schema Registry and SMM.

This improvement introduces the following changes in the affected connectors.

A new value, HWX Content-Encoded Schema Reference, is introduced for the Schema Access Strategy property

If this value is set, the schema is read from Schema Registry, and the connector expects that the Avro messages contain a content-encoded schema reference. That is, the message contains a schema reference that is encoded in the message content. The new value is introduced for the following connectors:

- ADLS Sink
- HDFS Sink
- HTTP Sink
- Influx DB Sink
- JDBC Sink
- JDBC Source
- Kudu Sink
- S3 Sink

The Schema Write Strategy property is removed from the following connectors

- ADLS Sink
- HDFS Sink
- S3 Sink
- InfluxDB Sink

A new property, Avro Schema Write Strategy is introduced

This property specifies whether and how the record schema is attached to the output data file when the format of the output is Avro. The property supports the following values:

- Do Not Write Schema: neither the schema nor reference to the schema is attached to the output Avro messages.
- Embed Avro Schema: the schema is embedded in every output Avro message.
- HWX Content-Encoded Schema Reference: a reference to the schema (identified by Schema Name) within Schema Registry is encoded in the content of the outgoing Avro messages.

This property is introduced for the following connectors:

- ADLS Sink
- HDFS Sink
- S3 Sink
- SFTP Source
- Syslog TCP Source
- Syslog UDP Source



Note: With the exception of InfluxDB Sink, this property replaces Schema Write Strategy in connectors where Schema Write Strategy was previously available.

The minor or major version of all affected connectors is updated

Existing connectors will continue to function, upgrading them, however, is not possible. If you want to use the new version of the connector, you must deploy a new instance of the connector.

For more information, see the documentation for each connector in [Kafka Connectors in Runtime](#) and [Streams Messaging Reference](#).

What's New in Apache Knox

Learn about the new features of Knox customers in Cloudera Runtime 7.2.18:

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 7.2.18.

Auto-incrementing column

Kudu now supports backup and restore of tables with auto incrementing columns. The restored table will have the auto incrementing column values identical to the source table for every row. For more details, see [Non-unique primary key index](#).

Kudu Range-aware Data Placement

Kudu places new tablet replicas using an algorithm which is both range and table aware. This algorithm helps to avoid hotspotting that occurs if many replicas backing the same range are placed on the same few tablet servers. Hotspotting causes tablet servers to be overwhelmed with write or read requests and can result in increased latency for these requests. To avoid hotspotting, this algorithm avoids targeting the same set of tablet servers for a set of replicas created in parallel. Rather, it spreads the replicas across multiple tablet servers. For more information, see [How Range-aware replica placement in Kudu works](#).

Kudu multi-master config change

You can now remove or decommission the unwanted master role instances through Cloudera Manager. Also, you can recommission any decommissioned master role instance in a multi-master deployment. For more information, see [Removing Kudu masters through Cloudera Manager](#).

Improvements

None.

What's New in Apache Livy

Learn about the new features of Livy in Cloudera Runtime 7.2.18.

High Availability support added for Livy

Livy now supports high availability. If there are more than one Livy Server in the cluster, high availability is automatically enabled.

What's New in Apache Oozie

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

Spark 3 support in Oozie

Oozie introduced the new Spark 3 based Spark 3 actions. For more information, see [Spark 3 support in Oozie](#).

Make hive-site.xml, hbase-site.xml and sqoop-site.xml available for all Oozie actions

Now Oozie automatically copies the hive-site.xml, hbase-site.xml, and sqoop-site.xml to all action's classpath. For more information, see [Oozie and client configurations](#).

Handle Sqoop Teradata Connector parcels installation and configuration for Oozie

When you install a Sqoop Teradata connector parcel, Cloudera Manager will automatically make the necessary Jars and configuration available to Oozie's Sqoop action.

Database connection properties enhancements

When it comes to configuring database connections, simply providing a hostname, port, username, and password may not be sufficient. In order to optimize Oozie's database connection, you might need to manually construct lengthy connection and configuration strings using safety-valve settings. To simplify this process and enable finer control over Oozie's database connection, you can use several enhancements. For more information, see [Fine-tuning Oozie's database connection](#).

What's New in Apache Phoenix

Learn about the new features of Phoenix in Cloudera Runtime 7.2.18.

Phoenix now supports alternate HBase connection registries

You can now use the Phoenix thick client using additional HBase connection registries. Phoenix now supports a Zookeeper-less connection strategy using a Master Registry implementation.

Phoenix now introduces the following JDBC URL variants.

- jdbc:phoenix+zk: Uses Zookeeper. This is the original registry supported since the inception of HBase and Phoenix.
- jdbc:phoenix+rpc: Uses RPC to connect to the specified HBase Region Server or Master nodes.
- jdbc:phoenix+master: Uses RPC to connect to the specified HBase Master nodes.

For more information, see [Using the Phoenix JDBC Driver](#).

What's New in Apache Ranger

The following new features and enhancements are generally available for Ranger customers in Cloudera Runtime 7.2.18:

Ranger Usersync option to update group memberships when same users and groups are synced from multiple sync sources

Ranger Usersync now provides an option for customers to treat users/groups from multiple sync sources as the same for updating group memberships. For more information, see the updated topic: [Configuring Usersync to sync directly with LDAP/AD](#).

HA support for Ranger Tag Sync/User Sync

Ranger now supports high availability for Ranger Tag Sync/User Sync. Configuring high availability adds another instance of each role to an additional host, which host continues to run the features if the default host fails. .

New Ranger API to collect metrics in Ranger Admin

Ranger now provides two APIs to fetch ranger admin metrics. One returns a response in JSON format and the other returns a response in prometheus-compatible format. For more information, see [Ranger Admin Metrics API](#).

New Ranger APIs to import/export roles in Ranger Admin

Ranger now includes APIs to import and export roles. For more information, see [Ranger REST API documentation](#).

Add support for enabling audit file accumulation

You can enable and configure alerts for Ranger plugin-supported services through Cloudera Manager. Such alerts notify when audit spool files accumulate in the spool directories for Solr and HDFS. For more information, see [Configuring audit spool alert notifications](#).

Add support for additional methods in RangerKafkaAuthorizer

RangerKafkaAuthorizer includes ACL APIs that refer to Ranger Policies when these commands are executed. Ranger relies on the grant, revoke and policy engine APIs to cater the needed functionality. For more information, see [Kafka ACL APIs support](#).

Add APIs to support force deletes of external users and groups from Ranger db

A Ranger database may (over-)populate with user and group records. To aid in removal of unnecessary users/groups, customers may use this feature to delete specific external user/groups or even all external users/groups if required. For more information, see [Force deletion of external users and groups from the Ranger database](#).

Ranger RMS support for s3 (Preview)

In CDP 7.2.18, Ranger RMS will support authorization for s3 storage locations, when deployed in an AWS environment. RMS for s3 will provide authorization for both HDFS and s3 file systems. A customer with this new RMS entitlement `ENABLE_RMS_ON_DATA LAKE` should be able to create a cluster with RMS as a configurable option (`--enable-ranger-rms`) through a `cdp cli` command `create-aws-datalake`. When RMS is selected during cluster setup, customers will not be required to install & configure RMS separately. For more information, see the updated topics and examples throughout [Ranger RMS - HIVE-S3 ACL Sync Overview](#).



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

What's New in Schema Registry

Learn about the new features of Schema Registry in Cloudera Runtime 7.2.18.

AvroConverter support for KConnect logical types

AvroConverter now converts between Connect and Avro temporal and decimal types.

Support for alternative jersey connectors in SchemaRegistryClient

`connector.provider.class` can be configured in Schema Registry Client. If it is configured, `schema.registry.client.retry.policy` should also be configured to be different than default.

This also fixes the issue with some third party load balancers where the client is expected to follow redirects and authenticate while doing that.

Upgraded Avro version to 1.11.1

Avro got upgraded from version 1.9.1 to 1.11.1.

New fingerprint version is added to Schema Registry with configuring option

A new fingerprint version, V2 is available in Schema Registry that contains the missing schema parts from the previous version. Newly created 7.2.18 clusters use the V2 fingerprint version. Upgraded clusters still use the V1 fingerprint version, but the `schema.registry.fingerprint.version` property can be used to change the fingerprint version in Schema Registry. Cloudera recommends to change the fingerprint version to V2 after upgrading to 7.2.18.

Support for additional JVM options

Additional JVM options can be passed to Schema Registry using the `schema.registry.additional.java.options` property in Cloudera Manager.

What's new in Apache Solr

Learn about the new features of Apache Solr in Cloudera Runtime 7.2.18.

- ZooKeeper SSL can now be configured for Solr and HBase Indexer. For more information, see [Enabling ZooKeeper SSL/TLS for Solr and HBase Indexer](#).
- Apache Solr is updated from 8.4.1 to 8.11.2 in this release of Cloudera Runtime. For more information, see [Apache Solr Release Notes](#) in the upstream documentation. For the list of notable unsupported features, see [Unsupported features](#).
- Solr now supports rolling upgrades. This means that rolling upgrades are available from release 7.2.18 to higher. Upgrades to release 7.2.18 still involve service downtime.
- Using Local File System (LFS) for both MapReduce Indexer Tool (MRIT) and HBase MRIT is now supported.
- Spark 3 is now supported.

- Spark 2 is deprecated in this release and support will be dropped in an upcoming release.
- This release introduces the following two health checks for the Solr service which give information about the status of the cores hosted on different hosts:

Recovering cores

By default this check reports concerning health if any of the hosted cores are in recovering status. This threshold can be modified in the configurations with the `solr_recovering_core_thresholds` configuration parameter.

Critical cores

By default this check reports "bad health" if any of the hosted cores are in down or recovery failed status. This threshold can be modified in the configurations with the `solr_critical_core_thresholds` config.

These checks are enabled by default for the Infra Solr service but disabled by default for the Workload Solr services (Cloudera Search).

- Critical CVE fixes.

What's New in Apache Spark

Learn about the new features of Spark in Cloudera Runtime 7.2.18.

Spark 3 support in Oozie

Oozie introduced the new Spark 3 based Spark 3 actions. For more information, see [Spark 3 support in Oozie](#)

Spark 3.4 support

Spark 3.4 is now supported in 7.2.18.

Spark History Server with High Availability

You can configure the load balancer for Spark History Server (SHS) to ensure high availability, so that users can access and use the Spark History Server UI without any disruption. Learn how you can configure the load balancer for SHS and the limitations associated with it. For more information, see [Using Spark History Servers with high availability](#)

Spark cluster template update

The following Spark 2 templates were deleted:

- Data Engineering: Apache Spark, Apache Hive, Apache Oozie
- Data Engineering: HA: Apache Spark, Apache Hive, Apache Oozie
- Real-time Data Mart: Apache Impala, Hue, Apache Kudu, Apache Spark
- Data Discovery and Exploration

The following Spark 3 templates were added:

- Data Discovery and Exploration for Spark3
- Data Engineering: Apache Spark3 cluster template was renamed to Data Engineering: Apache Spark3, Apache Hive, Apache Oozie

Spark 3 versions which are equivalent to the deleted Spark 2 templates mentioned above:

- Data Engineering: HA: Apache Spark3, Apache Hive, Apache Oozie
- Data Engineering: Apache Spark3, Apache Hive, Apache Oozie
- Real-time Data Mart: Apache Impala, Hue, Apache Kudu, Apache Spark3
- Data Discovery and Exploration for Spark3

For more information, see [Data Engineering clusters](#), [Data Mart clusters](#), and [Data Discovery and Exploration clusters](#).

Support for fault tolerant Spark Atlas hook

The `spark.lineage.kafka.fault-tolerant.timeout.ms` parameter was added to configure the Spark Atlas Connector so that Spark jobs can run even when Kafka brokers are down. This ensures that your job submissions do not fail. For more information, see [Spark connector configuration in Apache Atlas](#).

Spark 2 deprecation

Spark 2 was deprecated as of 7.2.17. See , [Deprecation Notices for Spark 2](#) .

What's New in Sqoop

Learn what's new in the Apache Sqoop client in Cloudera Runtime 7.2.18.

To access the latest Sqoop documentation on Cloudera's documentation web site, go to [Sqoop Documentation 1.4.7.1.6.0](#).

Secure options to provide Hive password during a Sqoop import

When importing data into Hive using Sqoop and if LDAP authentication is enabled for Hive, the necessity to set the Hive password parameter directly in the command-line poses a potential vulnerability. Passwords provided in plaintext within command-line interfaces are susceptible to unauthorized access or interception, compromising sensitive credentials and, subsequently, the security of the entire data transfer process.

Learn about the secure options that you can use to provide the Hive password during Sqoop-Hive imports instead of the earlier way of providing the password as plaintext in the command-line interface. For more information, see [Secure options to provide Hive password during a Sqoop import](#)

Sqoop Teradata Connector support for ORC file format

A new version of Cloudera Connector Powered by Teradata version 1.8.5.1c7 is released which includes ORC support in the Sqoop-Connector-Teradata component. You can use Teradata Manager to import data from the Teradata server to Hive in ORC format. For more information, see [Cloudera Connector Powered by Teradata Release Notes](#)

Discontinued maintenance of direct mode

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

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

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

What's new in Streams Messaging Manager


Learn about the new features of Streams Messaging Manager in Cloudera Runtime 7.2.18.

UI updates

Various improvements and new features are introduced for the SMM UI. The notable changes are as follows:

Cruise Control UI

A new page is added to Streams Messaging Manager to monitor the Kafka cluster state and rebalancing process with Cruise Control. The Cruise Control User Interface (UI) enables








you to review and configure the rebalancing of Kafka clusters through dashboards and a rebalancing wizard. The available goals and anomaly detectors are based on the Cloudera Manager configurations of Cruise Control. You can access Cruise Control from SMM using the  on the navigation sidebar.

For more information about Cruise Control in SMM, see [Monitoring and managing Kafka cluster rebalancing](#).

Data Explorer

- When you view Avro data in the Data Explorer, logicalTypes are converted by default. That is, instead of showing the underlying type, (for example, byte) the Data Explorer displays proper deserialized values.
- Avro messages are now pretty printed when you open them using the Show More option.

Kafka Connect

- Hovering over the status icons of connector tasks now displays the status text instead of the name of the icon.
- The Add missing configurations option now populates missing properties with default values.
- Adding flow.snapshot into a key field of a password type property clears password placeholders.
- The value field of the flow.snapshot property is now always a text area. Previously, if the property was added manually, the value field was a text field instead of an area.
- Text found in the  (Help) tooltip of connector property values now displays properly. Long strings no longer overflow the tooltip. Additionally, property descriptions are truncated. Clicking  or more... displays a pop-up containing the full description.
- Validating a connector configuration when the Kafka service is stopped returns a Connection refused error instead of validation passing.
- Deploying a new connector that has the same name as an existing connector no longer updates the existing connector. Instead, the connector deployment fails with Connector [***NAME***] already exists.
- A loading animation is displayed when loading connector templates.
- The horizontal divider found in the  context menu of properties is no longer displayed if the  String,  Boolean,  Number, and  Password options are not available for the property

For more information regarding the various new features and options related to Kafka Connect, see [Managing and monitoring Kafka Connect using Streams Messaging Manager](#).

Other

All charts present on the UI received a visual update. Additionally, more details are presented about the data they display.

Changes in Prometheus setup and configuration

Kafka Connect is now capable of securing its metric reporter with TLS/SSL and Basic Authentication. As a result, the setup steps required to configure Prometheus as the metrics store for SMM are changed. For updated deployment instructions, see [Setting up Prometheus for Streams Messaging Manager](#).

SMM internal Kafka topics are created with a replication factor of 3

From now on the __smm* internal SMM topics are created with a replication factor of 3. This change is only true for newly deployed clusters. The replication factor is not updated during the upgrade. Cloudera recommends that you increase the replication factor of these topics to 3 with kafka-reassign-partitions following an upgrade.

Remove keystore from SMM Schema Registry client configuration if Kerberos is enabled for Schema Registry

SMM uses a Schema Registry client to fetch schemas from Schema Registry. This Schema Registry client has Kerberos authentication properties and keystore properties for mTLS. Typically, the Schema Registry server, by default, does not allow mTLS authentication. But if mTLS is enabled in the Schema Registry server, then mTLS authentication has a higher precedence than Kerberos. Therefore, the mTLS principal (from the keystore) is used for authorization with Ranger rather than the Kerberos principal. This might result in authorization failures if the mTLS principal is not added to Ranger to access the Schema Registry resources.

From now on, the Schema Registry client used by SMM does not have keystore properties for mTLS when Kerberos is enabled. As a result, even if mTLS is enabled for the Schema Registry server, the Kerberos principal is used for authentication and authorization with Ranger.

Dedicated endpoint for connector creation

A dedicated endpoint for connector creation is introduced. The endpoint is `POST /api/v1/admin/kafka-connect/connectors`. Requests made to this endpoint fail with the following error message if the connector name specified in the request exists.

```
{
  "error_code": 409,
  "message": "Connector [***NAME***] already exists"
}
```

In previous versions, `PUT /api/v1/admin/kafka-connect/connectors/{connector}` was the only endpoint you could use to create connectors. However, this endpoint is also used to update connectors. If you specify the name of an existing connector in the request, the endpoint updates existing connector. As a result, Cloudera recommends that you use the new endpoint for connector creation going forward. The SMM UI is also updated and uses the new endpoint for connector creation.

Jersey client timeout now configurable

SMM uses internal Jersey clients to make requests to Kafka Connect and Cruise Control. The connection and read timeouts for these clients was previously hard-coded to 30 seconds. Configuring them was not possible. This release introduces new properties, which enable you to configure the connection and read timeouts of these clients. The default timeout remains 30 seconds. The properties introduced are as follows.

- Kafka Connect Client Connect timeout
- Kafka Connect Client Read timeout
- Cruise Control Client Connect timeout
- Cruise Control Client Read timeout

What's New in Streams Replication Manager

Learn about the new features of Streams Replication Manager in Cloudera Runtime 7.2.18.

The `--to` option in `srm-control` now creates the file if it does not exist

From now on, `srm-control` creates the file specified with the `--to` option if the file does not exist.

Verification of internal metrics topics can now be disabled

A new property, Verify Partition Count Of The Metrics Topic is introduced for the SRM service. This property controls whether SRM verifies the partition count of the `srm-metrics-[***SOURCE CLUSTER ALIAS***].internal` topics (raw metric topics) when SRM is started. This property is selected by default. Cloudera recommends that you keep this property selected. During certain upgrades, the property is set to false automatically for the duration of the upgrade to avoid upgrade issues.

Prefixless replication with the IdentityReplicationPolicy

Full support, including replication monitoring, is introduced for the `IdentityReplicationPolicy`. Unlike the `DefaultReplicationPolicy`, this policy does not rename remote (replicated) topics on target clusters. That is, the topics that you replicate will have the same name in both source and target clusters. This replication policy is recommended for deployments where SRM is used to aggregate data from multiple streaming pipelines. Alternatively, this replication policy can also be used if the deployment requires MirrorMaker1 (MM1) compatible replication.

Prefixless replication is enabled in Cloudera Manager with the `Enable Prefixless Replication` property. This property configures SRM to use the `IdentityReplicationPolicy` and enables internal topic based remote topic discovery, which is required for replication monitoring.

Limitations:

- Replication loop detection is not supported. As a result, you must ensure that topics are **not** replicated in a loop between your source and target clusters.
- The `/v2/topic-metrics/{target}/{downstreamTopic}/{metric}` endpoint of SRM Service v2 API does not work properly with prefixless replication. Use the `/v2/topic-metrics/{source}/{target}/{upstreamTopic}/{metric}` endpoint instead.
- The replication metric graphs shown on the **Topic Details** page of the SMM UI do not work with prefixless replication.



Important: If you have been using the `IdentityReplicationPolicy` in a previous version of Cloudera Runtime, ensure that you transition your configuration and set the `IdentityReplicationPolicy` with `Enable Prefixless Replication`. If you do not transition your configuration, replication monitoring will not function.

For more information, see

- [Streams Replication Manager replication flows and replication policies](#)
- [Enabling prefixless replication](#)

Internal topic based remote topic discovery

From now on, SRM uses an internal Kafka topic to keep track of remote (replicated) topics. Previously, SRM relied on the naming conventions (prefixes) used by the `DefaultReplicationPolicy` to discover and track remote topics.

This feature enables SRM to provide better monitoring insights on replications. Additionally, if the feature is enabled, SRM is capable of providing replication monitoring even if a replication policy other than the `DefaultReplicationPolicy` is in use. Most notably, this enables replication monitoring when SRM is configured for prefixless replication with the `IdentityReplicationPolicy`.

This feature is enabled in Cloudera Manager by selecting the `Remote Topics Discovery With Internal Topic` property. The property is selected by default on newly deployed clusters, but must be enabled manually for existing clusters after an upgrade. Cloudera recommends that you enable this feature no matter what replication policy you are using.

For more information, see [Streams Replication Manager remote topic discovery](#).

Configurations to customize replication-records-lag metric calculation

Three new properties are introduced that enable you to control how SRM calculates the `replication-records-lag` metric. This metric provides information regarding the replication lag based on offsets. The metric is available both on the cluster and the topic level. The following new properties are introduced because the calculation of the metric with default configurations might add latency to replications and impact SRM performance. While these properties are configured in Cloudera Manager, they do not have dedicated configuration entries. Instead, you add them to Streams Replication Manager's Replication Configs to configure them.

Table 1:

Property	Default Value	Description
replication.records.lag.calc.enabled	true	Controls whether the replication-records-lag metric is calculated. This metric provides information regarding the replication lag based on offsets. The metric is available both on the cluster and the topic level. The calculation of this metric might add latency to replications and impact SRM performance. If you are experiencing performance issues, you can try setting this property to false to disable the calculation of replication-records-lag. Alternatively, you can try fine-tuning how SRM calculates replication-records-lag with the replication.records.lag.calc.period.ms and replication.records.lag.end.offset.timeout.ms properties.
replication.records.lag.calc.period.ms	0	Controls how frequently SRM calculates the replication-records-lag metric. The default value of 0 means that the metric is calculated continuously. Cloudera recommends configuring this property to 15000 ms (15 seconds) or higher if you are experiencing issues related to the calculation of replication-records-lag. A calculation frequency of 15 seconds or more results in the metric being available for consumption without any significant impact on SRM performance.
replication.records.lag.end.offset.timeout.ms	6000	Specifies the Kafka end offset timeout value used for calculating the replication-records-lag metric. Setting this property to a lower value than the default 6000 ms (1 minute) might reduce latency in calculating replication-records-lag, however, replication-records-lag calculation might fail. A value higher than the default can help avoid metric calculation failures, but might increase replication latency and lower SRM performance.

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 7.2.18.

Apache Hadoop YARN

Mixed resource allocation mode (Technical Preview)

You can now specify the resources in mixed types. You can specify the actual units of vcores and memory resources for each queue or specify the percentage of the total resources used by each queue or specify a weight for each queue. You can use a combination of these allocation modes. The queues under one parent can also mix their modes. For more information, see [Mixed resource allocation mode](#)

YARN Queue Manager

Java heap size can now be configured

You can now customize Java heap size in YARN Queue Manager. Although the default for this setting should be valid in most deployment scenarios, you have the option to update the setting only if a given cluster has run into memory-management issues, otherwise, the settings can remain.

What's New in Apache ZooKeeper

Learn about the new features of ZooKeeper in Cloudera Runtime 7.2.18.

Rebase ZooKeeper and Curator

CDP Public Cloud is updated to use Apache ZooKeeper version 3.8.1 and Apache Curator version 5.4.0 for a smoother and better functionality. Upgrade your client applications for seamless connectivity.

ZooKeeper client configuration changes

Cloudera Manager now has the capability to dynamically generate ZooKeeper client configuration under `/etc/zookeeper` based on service settings.

The `zookeeper-client` script has been enhanced to take advantage of SSL/SASL settings.

Unaffected Components in this release

There are no new features for the following components in Cloudera Runtime 7.2.18.

- Data Analytics Studio
- Apache Hadoop HDFS

Fixed Issues In Cloudera Runtime 7.2.18

You can review the list of reported issues and their fixes in Cloudera Runtime 7.2.18.

Fixed Issues in Atlas

Review the list of Apache Atlas issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-65258: Iceberg table does not show lineage in Spark for dataframe

Added Dataframe / AtomicReplaceTableAsSelectExec to Spark Atlas Connector for Spark3 which is required for Iceberg.

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

Added CTAS Iceberg support to Spark Atlas Connector for Spark3.

CDPD-65199: Post upgrade (both Data Lake and Data Hub) to swap OS and dl shape together from Cloudera Runtime 7.2.15 to 7.2.18, one of the Atlas server throws "Could not retrieve active server address as it is null"

Fixed the way ZNode creation is done. Added a check to see if the Znode is created from the same session or not.

CDPD-65087: Iceberg table entities not getting reflected in Atlas after upgrade from Cloudera Runtime 7.2.17 to 7.2.18

After upgrading from 7.2.17 to 7.2.18 for the Iceberg type entities, all details were not visible on the entity details page due to changes in the type definition of Iceberg table and column. This patch resolves the issue.

CDPD-64125: Atlas server side Ignore and Prune patterns does not function as expected

Fixed the bug where ignore and prune pattern was not working on Atlas server side.

CDPD-61587: Stored cross-site scripting on "Description" field under classification

Sanitized the value of the description field against XSS script passed through UI and Curl command before rendering on the UI.

CDPD-57862: Atlas - Upgrade grpc to 1.53.0+ due to CVE-2023-32732 and CVE-2023-1428

Upgrade grpc to 1.53.0+ due to CVE-2023-32732 and CVE-2023-1428.

CDPD-58592: [CKP4 (same value)] Atlas not null filter on classification returns null values

Cause: It is because of Solr version upgrade, until 8.4.1, Solr supported non empty string. Fix: For IndexQuery : ["" TO *] works to get nonEmpty field entities. For Inmemory Predicates: Used NonEmptyPredicate.

CDPD-58492: Atlas - Upgrade Netty Project to 4.1.94. Final due CVE-2023-34462

Upgrade Netty Project to 4.1.94.Final from 4.1.86.Final.

CDPD-57453: Atlas Error while writing audits to GCP Data Lake

Removed the lib which was causing the conflict.

CDPD-57433: Atlas - Upgrade gremlin shaded to 3.5.5+ due to jackson-databind CVEs

Upgrade gremlin shaded to 3.5.5 from 3.5.4.

CDPD-55876: Atlas - Upgrade Spring Security to 5.7.8+/5.8.3+/6.0.3+ due to CVE-2023-20862

Upgrade Spring Security to 5.8.3 to 5.7.5.

CDPD-55858: [Atlas] Relationship search isNull/notNull case does not consider null strings

It is fixed by CDPD-58592.

CDPD-55769: Atlas - Upgrade Spring Framework to 5.3.27/6.0.8 due to CVE-2023-20861, CVE-2023-20860 and CVE-2023-20863

Upgrade Spring Framework to 5.3.26/6.0.7 due to CVE-2023-20861.

CDPD-55617: Atlas - Upgrade Nimbus-JOSE-JWT to 9.24 due to CVEs coming from json-smart

Upgrade Nimbus-JOSE-JWT to 9.24 from 9.8.1.

CDPD-55440: Atlas - Upgrade snakeyaml to 2.0 due to CVE-2022-1471

Upgrade snakeyaml to 2.0 from 1.33.

CDPD-55252: Atlas - Upgrade jackson-databind to 2.12.7.1/2.13.4.1+ due to CVE-2022-42003, CVE-2022-42004

Upgrade jackson-databind to 2.12.7.1.

CDPD-54846: Atlas: CVE-2023-24998-upgrade commons-fileupload library to version 1.5

Upgrade commons-fileupload library to version 1.5 from 1.3.3.

CDPD-54645: Ranger tag sync for Iceberg table type

Added Iceberg support for Ranger Tagsync.

CDPD-54391: Zookeeper SSL/TLS support for Ranger

Ranger Admin, Tagsync, Usersync, RAZ, and RMS supports SSL/TLS secure connection with Zookeeper.

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-50913: Atlas - Upgrade azure-storage libraries due to CVE-2022-30187

Azure Storage Library Information Disclosure Vulnerability.

CDPD-50416: UI: Setting 'atlas.ui.date.format' to certain value creates incorrect date entries

Correct date format is displayed after setting the default date format through application properties file.

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-47524: [PbC] Sync Atlas code with Apache

Atlas code base needs to be constantly in sync with Apache.

CDPD-45073: Implement aging for audits stored by Atlas.

Feature to reduce Atlas audit storage by aging out audit data, using different criteria like TTL and audit count.

Apache patch information

- ATLAS-4805
- ATLAS-4768
- ATLAS-4762
- ATLAS-4750
- ATLAS-4732

Fixed Issues in Avro

There are no fixed issues for Avro in Cloudera Runtime 7.2.18.

Apache patch information

None

Fixed Issues in Cloud Connectors

Review the list of Cloud Connectors issues that are resolved in Cloudera Runtime 7.2.18.

- CDPD-62811: Hadoop ABFS driver fails with HTTP 429 throttling issues
- CDPD-47679: Add authorization check for x-amz-copy-source for users in RAZ

Apache Patch Information

- HADOOP-17377
- HADOOP-17386
- HADOOP-18233
- HADOOP-18688
- HADOOP-18705
- HADOOP-18724
- HADOOP-18752
- HADOOP-18757
- HADOOP-18781
- HADOOP-18793
- HADOOP-18845
- HADOOP-18873
- HADOOP-18915
- HADOOP-18925
- HADOOP-18940
- HADOOP-18997
- HADOOP-19015
- HADOOP-19027

- HADOOP-19033
- HADOOP-19046
- MAPREDUCE-7432
- MAPREDUCE-7435

Fixed issues in Cruise Control

Review the list of Cruise Control issues that are resolved in Cloudera Runtime 7.2.18.

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

This issue has been fixed as the Cloudera Manager Metrics Reporter is deprecated from Cruise Control.

OPSAPS-68148: Cruise Control rack aware goal upgrade handler

As Cruise Control automatically overrides all occurrences of the deprecated RackAwareGoal with RackAwareDistributionGoal during upgrade, the customized values of the Cruise Control goals will remain the same and there is no need to manually provide the values after an upgrade.

Fixed Issues in Apache Hadoop

There are no fixed issues for Hadoop in Cloudera Runtime 7.2.18.

Apache Patch Information

None

Fixed Issues in HBase

Review the list of HBase issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-58799: 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-54370: Implement calling LeaseRecovery in HBase client

Use LeaseRecoverable and SafeMode introduced in hadoop-common. This is similar to HBASE-27769 as part of the Cloudera HBase release.

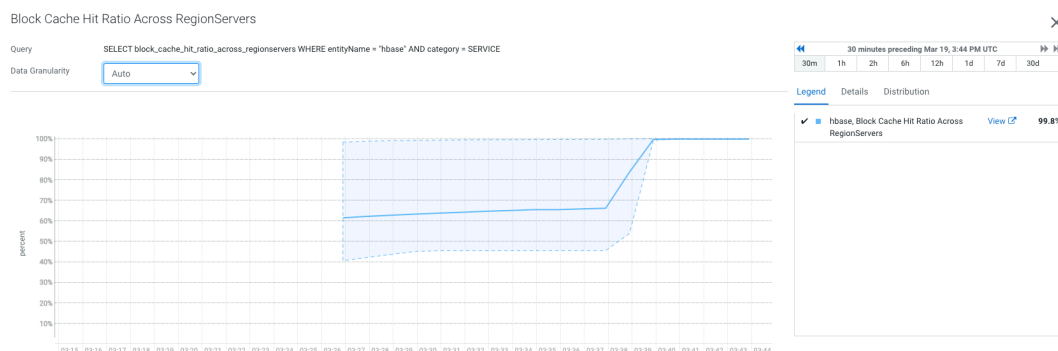
CDPD-47423: HBase - Support for JDK17 in all sub-components

JDK 17 support has been added

HBASE-28189: HBase supports accurate hit ratio report on Cloudera Manager

When you use Cloudera Operational Database (COD) over S3 with ephemeral cache, the HBase block cache uses CombinedBlockCache for implementation along with a LruBlockCache instance for the metadata blocks, and a BucketCache using the ioengine file on the nvme SSD disk for the data blocks. An important metric to measure the efficiency of the cache is the *hit ratio*, which reports the frequency that the client reads have hit the cache to fetch its required data, rather than going to S3, which is much slower. Cloudera Manager provides hit ratio metrics-based charts, however the chart was inaccurate because the calculation of hit ratio metrics in the HBase code while using the CombinedBlockCache implementation was inaccurate.

This issue is now fixed. The following example of a Cloudera Manager chart shows an optimal ephemeral cache with close to 100% hit ratio across all RegionServers.



Apache Patch Information

- HBASE-27820
- HBASE-26038

Fixed Issues in HDFS

There are no fixed issues for HDFS in Cloudera Runtime 7.2.18.

Apache Patch Information

None

Fixed Issues in Apache Hive

Review the list of Hive issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-35164: Support creation of column histogram statistics

Hive supports the creation of column histogram statistics and uses it at query planning.

CDPD-49145: Oozie and Spark tests are failing in pre-phase (mul-comp-pre) with Zookeeper-based or direct JDBC URL to Hive

HiveServer uses `InetAddress.getHostName()` API to get its hostname and register itself with Zookeeper. The API behaviour changed on JDK 11 with specific operating systems to return just the hostname without the domain suffix. Therefore, HiveServer is inaccessible to clients when the server information is obtained from Zookeeper. To address this issue, the `InetAddress.getCanonicalHostName()` API is used to fetch the hostname with the fully qualified domain.

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 patch.

CDPD-50730: Hive WebUI HTTP 500 error due to JAR order in classpath

The issue was fixed by removing `javax.servlet.jsp-api` dependency from HiveServer2, which helps to avoid the intermittent `NullPointerException` while opening the home web page.

CDPD-51885: Column statistics are not getting published after an insert query into an external table with custom location

The fix ensures that the column status is published after an insert into the external table created with an empty custom location.

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

Few datatypes are not mapped from Postgres or Redshift to Hive data types in the connector, which resulted in displaying null values for the columns of those data types. This issue is fixed.

CDPD-56205: [Security Vulnerability reported] added redirect url check in strategy

Added a validation to ensure the URL used during SSO workflow is proper (http/https).

CDPD-57988: Modify the conditions to enter the BETWEEN histogram selectivity code and account for NULLs

Fixed incorrect filter selectivity of BETWEEN expressions when using histograms.

CDPD-63449: Alter table add partition is failing with error: "java.io.IOException: Got exception: java.nio.file.AccessDeniedException"

Alter table add partition runs without any exception. The issue was noticed on the HMS API side and was fixed.

CDPD-64293: Compaction cleaner is turned off by default in Public Cloud resulting in compaction test failures

The value of the hive.compactor.cleaner.on property is set to 'true' for the compactor to run on the HMS instance.

Apache Patch Information

- HIVE-27116
- HIVE-27147
- HIVE-27163
- HIVE-27179
- HIVE-27316
- HIVE-27554

Fixed Issues in Hive Warehouse Connector

Review the list of Hive Warehouse Connector (HWC) issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-60192: Unable to load data with array type through Hive Warehouse Connector

The fix addresses an issue where users were unable to load data with array type through Hive Warehouse Connector

CDPD-62507: HWC - Upgrade netty to 4.1.100.Final due to CVE-2023-44487

In HWC, netty was upgraded from 4.1.86.Final to 4.1.100.Final due to CVE-2023-44487

Fixed Issues in Hue

Review the list of Hue issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-59802: Hue is unable to process parallel requests

Due to an issue with Impala file reader not getting closed after use, it caused a leak in the connection pool. Because of the leak, the Hue Query Processor could not read Impala event files after a few queries. This issue has been fixed by ensuring that the file reader is closed, and ingestion of Impala query works as expected.

CDPD-45438: Hue Importer does not validate the column name; shows an error if the column name contains a space

If column names contained spaces in the file you were trying to import into Hue, then you encountered the following error: AnalysisException: Invalid column/field name. When you fixed the column name and proceeded with importing the file, the rename operation also failed. This issue has been fixed. Hue now validates the column names to verify whether they contain spaces and does not allow you to proceed without correcting the column names.



Note: Hue only allows alphanumeric characters and underscores (_) in column names for all dialects and engines.

TSB 2023-703: Risk of data loss when using Hue S3 File Browser

This issue has been fixed.

CDPD-54714: SSO does not work while logging in from the Hue UI

Earlier, due to a missing configuration in Cloudera Manager, SSO did not work when you enabled Knox as an authentication backend and when Hue was in HA mode. This issue has been fixed.

CDPD-61550: ABFS File Browser lists a maximum of 5000 objects

Earlier, when you browsed ADLS Gen 2 storage using Hue's ABFS File Browser, it displayed only 5000 objects. This issue has been fixed by including a continuation token in the response header (x-ms-continuation).

CDPD-56024: The last status of a workflow task does not reset in Hue

Earlier, when you reran an Oozie job or workflow from the Hue web interface, the task status continued to display the old status and did not reset. This has been fixed.

CDPD-59896: Hue fails to clean up documents if Oozie is on the blocked list of services

Earlier, when you ran the "desktop_document_cleanup" script to clean old documents, and if you added Oozie to the blocked list of apps in the Hue Advanced Configuration Snippet, the script failed. To run the script successfully, you had to remove the Oozie service from the app_blacklist property manually. This issue has been resolved. You no longer need to manually remove Oozie from the blocked list of apps.

CDPD-61589: Downloading a file from ABFS using Hue returns a corrupted file

Earlier, downloading a file from ABFS using the ABFS File Browser returned a corrupted file because of reading 1 extra byte for every 1 MB chunk size. This issue has been fixed.

CDPD-50554: Unable to deselect the checkbox on Hue's Job Browser page after a transition

Earlier, when you selected a job from the Running list from the Job Browser page and terminated it, the job remained selected after transitioning to the Completed list. This issue has been fixed.

CDPD-50493: Sample data from Table Browser in Hue launches expensive queries from the Impala Views

Earlier, when you tried to view sample data from the Table Browser in Hue, Hue automatically ran the following query: "SELECT * FROM <database>.<table> LIMIT 100" from tables and views. To generate data from views, Hue fetched all rows and then displayed 100, based on the limit. This issue has been fixed. Hue no longer fetches any sample data from views.

CDPD-57894: Unable to filter Oozie tasks based on status in Hue

The ability to filter Oozie tasks based on their status from the Job Browser Schedules page that was removed has now been added.

CDPD-61651: Schema and tables are not displayed in the Phoenix Editor as Phoenix is case-sensitive

Schema and table names in Phoenix are case-sensitive. Earlier, Hue used the UPPER() function while processing queries, resulting in issues when running Phoenix queries. This issue has been resolved by removing the UPPER() function for databases and tables.

CDPD-61606: Potential data loss while moving files or folders using the Hue S3 File Browser

When you attempted to copy or move files or folders using the Hue File Browser, you might have encountered data loss if you did not wait for the **Move to** modal to load completely before clicking Move, or if the source and destination paths were the same. This issue has been fixed.

CDPD-58714: Cannot rerun multiple jobs (scheduled tasks of a coordinator job)

Earlier, when you selected more than one job on the Jobs Schedule page in Hue, the Rerun button was disabled. You could only run one job. This issue has been fixed. You can now rerun multiple jobs in Hue.

CDPD-57919: YARN does not send the active ResourceManager to Hue

If there is a ResourceManager failover, then Hue continues to send requests to the old ResourceManager instance (which is on standby mode). When you tried to terminate the job, the kill

job command failed with the following error: YARN RM returned a failed response: This is standby RM. The redirect url is: /ws/v1/cluster/apps/application_1687323161493_0002/state?doAs=admin (error 307). You had to restart Hue to temporarily resolve the issue. This issue has been fixed.

CDPD-58012: Spark SQL editor (Spark3 with Livy) does not reflect output when rendering array and map columns

Earlier, the Spark SQL editor in Hue did not display column results for array, map, and struct complex data types. The same query that included these data types worked fine from Hive and Impala editors and showed results as expected for the concerned columns. This issue has been fixed.

CDPD-48059: Unable to view and access S3 buckets from Hue configured with RAZ

Earlier, when you configured access to AWS S3 buckets using RAZ in the following regions, you could not access S3 buckets from the Hue File Browser: cn-*, eu-central, ap-northeast-2, ap-south-1, us-east-2, ca-central and eu-west. Due to a bug in AWS Boto SDK, accessing S3 buckets from Hue using RAZ was blocked. This issue has been fixed.

CDPD-55430: Debug-level logging is missing in Hue server logs

The Hue server logs contained only INFO-level information and DEBUG-level information was missing. Cloudera has fixed multiple logging issues, overall. Now, only INFO-level logging is on by default. You can turn on DEBUG-level logging from Cloudera Manager by selecting the Enable Django Debug Mode option in Hue configurations.

CDPD-60510: Query Processor fails to start on a FedRAMP cluster

Earlier, the Query Processor service did not start on a FedRAMP cluster, and you may have seen the following error in the logs: “ERROR io.dropwizard.cli.ServerCommand - Unable to start server, shutting down java.io.IOException: Invalid keystore format”. This happened because in FIPS mode, the Query Processor service expects to find its truststore in the Java KeyStore (JKS) format, even though it might be using the Bouncy Castle library (which supports the bcfsks format). This issue has been fixed by adding the keystore_type configuration (Java Keystore Type field) in Cloudera Manager through which you can specify the appropriate JKS type (in this case bcfsks), which enables the FIPS mode to work.

Fixed Issues in Apache Impala

Review the list of Impala issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-64905: Backport IMPALA-12589 to active branches

Fix NPE when querying external tables that points to a single file.

CDPD-60469: Impala log rotation not working on the old pid log files.

Allow automatic removal of old logs from previous PID.

CDPD-56871: Backport fix for IMPALA-12114 to impacted releases

An issue where idle Impala clients using TLS were needlessly disconnected has been fixed.

CDPD-55490: Impala - Upgrade Jetty to 9.4.51/10.0.14+/11.0.14+ due to CVE-2023-26048 and CVE-2023-26049

Transient imports via Hadoop/Hive/Ranger have been upgraded.

CDPD-55460: Impala - Upgrade Spring Framework to 5.3.27/6.0.8 due to CVE-2023-20863

Spring Framework has been upgraded to 5.3.27.

CDPD-41064: IMPALA-11360 Support Java11 in Impala

Impala supports Java 11.

Apache Patch Information

- IMPALA-12595
- IMPALA-12589

- IMPALA-12214
- IMPALA-12114

Fixed Issues in Apache Kafka

Review the list of Apache Kafka issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-62059: AvroConnectTranslator should handle null values in fromConnectData method

Fix possible NPE exception issues in connector tasks which operate with Avro data format using the Cloudera AvroConverter.

Apache patch information

None

Fixed Issues in Apache Knox

Review the list of Knox issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-63593: Graviton Support for Knox-Gateway

Added support for multi-arch builds

CDPD-62767: KnoxShell fails with Unsupported class file major version 61 error

Fixed an issue where on JDK 17 KnoxShell fails with error Unsupported class file major version 61 error.

CDPD-62057: DefaultDispatch doesn't forward inbound request headers in case of requestType=OPTIONS

Inbound request headers are forwarded with OPTIONS type HTTP requests.

CDPD-60911: Knox Readiness Awareness and Notification

Improved Knox readiness check

CDPD-54722: Accessing service via cdp-proxy-api failed with 404 with logs having SAXParseException

Fixed a race condition in the XML parser of Knox

CDPD-50773: Need a provision for deleting the custom topology/provider/descriptor

"remove" keyword can be used to remove custom descriptors and shared providers.

CDPD-49173: Knox - Support for JDK17 in all sub-components

Knox support for JDK 17

CDPD-48928: Rebase Knox to Apache 2.0.0

Knox in CDP was rebased on top of Apache Knox 2.0.0.

Apache patch information

- KNOX-2989
- KNOX-2960
- KNOX-2955
- KNOX-2923
- KNOX-2905
- KNOX-2899

Fixed Issues in Apache Kudu

Review the list of Apache Kudu issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-64527: Unable to place replicas using range aware logic with multiple locations

This patch fixes a bug that caused the master to crash when attempting to place replica using range aware replica placement logic.

KUDU-3500: Optimization to skip operations timed out in the prepare queue

Enhances tablet server efficiency by immediately responding with a TimedOut error for write operations that expire in the prepare queue, thus improving load handling and adding a new metric to track such occurrences.

KUDU-3520: Fix file descriptor leak in data-at-rest encryption

Fixed a file descriptor leak in PosixEnv related to encryption setup failures, ensuring cleanup in error scenarios.

KUDU-3515: Resolved Unbounded Range Partition Dropping Issue

Fixed an incompatibility issue that prevented dropping unbounded range partitions in tables with combined hash and range partitioning upon upgrading to Kudu 1.17.0, ensuring seamless partition management.

KUDU-3495: Fix for handling upsert requests with outdated client schemas

This update resolves an issue where upsert requests from clients using outdated schemas could unintentionally reset newly added columns to default values by initializing the tablet_isset_bitmap during request decoding.

KUDU-3489: Support large messages for Subprocess communication protocol

Kudu now supports messages of size up to 8MB by default to be transmitted between Kudu master and subprocess server. This is necessary to make Ranger authz policies work in a Kudu cluster with many tables.

KUDU-3461: Improved stability for Kudu clients with error handling for invalid tablet IDs

To prevent Impala daemon crashes due to stack overflow, Kudu C++ clients now detect invalid tablet ID conditions and return an error, enhancing system stability during concurrent operations.

KUDU-3459: Enhanced superblock download mechanism

Introduced a new feature, controlled by `--tablet_copy_support_download_superblock_in_batch`, to download superblock in segments, addressing issues with large superblocks exceeding the `--rpc_max_message_size` limit.

KUDU-3455: Enhanced hash partition pruning efficiency in clients

This enhancement reduces memory usage and accelerates hash partition pruning for in-list predicates in both Kudu C++ and Java clients, offering performance improvements with minimal memory requirements.

KUDU-3448: Secure storage for IPKI and TSK key material

Enhances security by encrypting IPKI and TSK key material on disk using a password-derived symmetric key, allowing for secure integration with hardware security modules or other password provision methods without storing the password on disk.

KUDU-3402: Update trace-viewer

Kudu tracing.html now works with modern browser versions.

KUDU-3359: Allow multi-JAR classpaths for Ranger client

In order to let Ranger client write audit to HDFS, Hadoop client JARs need to be loaded. To make sure this is possible to do, this fix changes the behavior of the `ranger_jar_path` to allow colon-separated JAR classpaths to be passed to Java.

Apache Patch Information

- KUDU-3532
- KUDU-3520

- KUDU-3515
- KUDU-3500
- KUDU-3495
- KUDU-3489
- KUDU-3461
- KUDU-3459
- KUDU-3455
- KUDU-3448
- KUDU-3402
- KUDU-3359

Fixed Issues in Apache Livy

Review the list of Apache Livy issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-61564: Spark - Caused by: java.lang.NoClassDefFoundError: org/datanucleus/store/query/cache/QueryCompilationCache

Upgraded datanucleus-core dependency to 5.2.10

CDPD-55423: remove verbose output on Livy UI Error pages.

We have added a new Livy configuration property "livy.server.send-server-version". This can be set to "true" to send the server version in CM. By default this properties is set as "false".

CDPD-55116: Fix Spark vulnerability CVE-2023-22946

This fix is blacklisting "spark.submit.deployMode" and "spark.submit.proxyUser.allowCustomClasspathInClusterMode" spark configurations in Livy create session REST API. We have added a new Livy configuration "livy.server.session.allow-custom-classpath" to allow custom class path. In order to disable or rollback this fix, we can add "livy.server.session.allow-custom-classpath" as "true" in Livy configuration via the CM safety valve.

Apache Patch Information

- LIVY-975
- LIVY-974

Fixed Issues in Apache Oozie

Review the list of Oozie issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-44209 SqoopMain's printArgs masks Sqoop command line option if preceding one contains "password"

In YARN, there was a previous issue in Oozie where command-line arguments were masked incorrectly due to mistaken password detection. As a resolution, customers now have the option to utilize the "oozie.launcher.argumentMaskingExceptionList" configuration. This feature allows them to specify exceptions for password masking. For detailed information on how to use this configuration, please refer to the documentation in oozie-default.xml.

CDPD-46049 SSH action fails when 'oozie.action.ssh.http.command.post.options' property contains double quotes

The SSH action's callback mechanism failed with "Invalid content-type" error when capture-output was used in the action definition.

CDPD-50296 Improve Oozie's app state action checking

Enhanced Oozie's action state checking, to immediately query for running applications right after start-up.

CDPD-50915 Oozie shouldn't ignore hive-site.xml on host if no hive-site is on Spark share lib

Now, Oozie will not ignore a *-site.xml file in a Spark or Spark3 action if it was specified manually via the --files Spark opts.

CDPD-55101 Invocation of Main class completed Message is skipped when LauncherSecurityManager calls system exit

Oozie will print out the 'Invocation of Main class completed' message in the Launcher AM logs even if there is a security Exception.

CDPD-56724 Oozie web console is allowing access to list directories

From now on, directory listing through the Oozie web console is disabled.

CDPD-65049: HTTP security headers are missing from Oozie response

Oozie now returns the HSTS headers in all of its API and UI servlets

CDPD-64730: Oozie LauncherAM memory settings cannot be applied

Fixed the issue where the Oozie LauncherAM memory settings couldn't be applied

CDPD-64376: Oozie's Spark and Spark3 option parser does not respect Java arguments starting with '--'

Fixed the issue where Oozie incorrectly split the arguments of a Spark or Spark3 action where a special Java argument (for example --add-exports) was used inside another Spark argument

CDPD-64133: The Oozie client should be able to handle Java 11+ related parameters

The Oozie command line tool now handles Java 11+ related (for example --add-exports) arguments correctly

CDPD-63724: Add spark-sql-kafka to Oozie Spark/Spark3 share libs

The spark-sql-kafka library was added to the Spark/Spark3 ShareLibs to make sure Spark-Atlas Hooks are always working as expected

CDPD-63326: Fix CVE-2023-36877 Apache Oozie Spoofing Vulnerability

Oozie sanitizes custom filters, hence preventing the execution of vulnerable scripts

CDPD-58538: Oozie should upload and use the config files from sqoop-conf/managers.d when available

Previously, Oozie did not honor Sqoop's managers.d configurations and extra connector Jars from the lib folder, but now both are automatically available in Oozie's Sqoop action, allowing users to seamlessly utilize connectors like the Sqoop Teradata connector without the need for manual configuration updates or copying Jars to the Workflow's lib folder

CDPD-56936: Oozie's db cli tool does not honour custom connection properties

The Oozie DB CLI tool did not respect the "ConnectionProperties" property set by the user through the "oozie.service.JPAService.connection.properties" configuration in Oozie

CDPD-56724: Oozie web console is allowing access to list directories

From now directory listing via the Oozie web console is disabled

CDPD-55101: Invocation of Main class completed Message is skipped when LauncherSecurityManager calls system exit

Oozie will print out the 'Invocation of Main class completed' message in the Launcher AM logs even if there is a security Exception

CDPD-50915: Oozie shouldn't ignore hive-site.xml on host if no hive-site is on Spark share lib

Oozie now will not ignore a *-site.xml file in a Spark or Spark3 action if it was specified manually via the --files Spark opts

CDPD-50296: Improve Oozie's app state action checking

Enhanced Oozie's action state checking, to immediately query for running applications right after start-up

CDPD-47821: Add missing Sqoop Atlas notification jars to Sqoop share lib

Earlier, Atlas notification was non-functional in Oozie's Sqoop action due to missing Jars, but with the inclusion of those Jars in Oozie's Sqoop ShareLib, Atlas notifications are now expected to function correctly in Oozie's Sqoop action.

CDPD-46049: SSH action fails when 'oozie.action.ssh.http.command.post.options' property contains double quotes

The SSH action's callback mechanism failed with "Invalid content-type" error when capture-output was used in the action definition

CDPD-44209: SqoopMain's printArgs masks Sqoop command line option if preceding one contains "password"

In Yarn, there was a previous issue in Oozie where command-line arguments were masked incorrectly due to mistaken password detection. As a resolution, customers now have the option to utilize the "oozie.launcher.argumentMaskingExceptionList" configuration. This feature allows them to specify exceptions for password masking. For detailed information on how to use this configuration, please refer to the documentation in oozie-default.xml.

Apache patch information

- OOOIE-3718
- OOOIE-3716

Fixed Issues in Phoenix

Review the list of Phoenix issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-58269: Add option to CREATE TABLE to skip verification of HBase table

A new option NOVERIFY has been introduced for CREATE TABLE command, that allows skipping the verification of columns with empty qualifier. This feature is useful when a Phoenix table is restored from HBase snapshot. It allows to skip the lengthy validation process when executing CREATE TABLE.

CDPD-55719: Cluster Launch with JDK 17 fails in first run command for OMID

JDK 17 support has been added

CDPD-49395: CR - ZooKeeperless Phoenix Client

Phoenix now supports alternate HBase connection registries. See https://phoenix.apache.org/classpath_and_url.html for details.

Apache Patch Information

- PHOENIX-6973
- PHOENIX-6523
- OMID-242

Fixed Issues in Parquet

There are no fixed issues for Parquet in Cloudera Runtime <version number>.

Apache Patch Information

None

Fixed Issues in Apache Ranger

Review the list of Ranger issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-65433: Execute and read permissions granted to a user in different HDFS policies does not take effect.

As part of this bug fix, execute and read permissions granted to a user in different HDFS policies are working as expected. For example:

Policy 1: Granted the "public" group "execute" permission to "/" HDFS policy recursively.

Policy 2: Granted only the "read" permission to the user for "/hdp"

Perform a list on "/hdp".

CDPD-65310: [7.2.18 CLONE] - Performance degradation while retrieving mapped hive resource for s3 location.

Retrieving mapped hive resource for s3 location will be faster.

CDPD-64800: Classic UI - Security zone form not populate resources value properly while creating and editing zone form.

After the patch, zone form populates resources value properly.

CDPD-63148: RAZ client should encode the authorization URL to support unicode characters

Call to RAZ authorization api to support non-english characters in the URL.

CDPD-62934: Insecure direct object reference

As part of this fix, audit metrics endpoint made secure.

CDPD-61584: [Intermittent] Active NN not getting latest resource mappings from RMS server

NameNode is HA-enabled and both NameNodes send requests to download deltas after the full-sync is performed in RMS; then both NN will get the latest resource mappings from the RMS server.

CDPD-60952: [7.2.18.0] - Add server side validation for service audit filter

As part of this fix added server-side validation for service audit filter.

CDPD-60870: Ranger KMS junit tests are failing

Unsupported cipher removed from UT.

CDPD-57635: 7.2.18 -pre-cdpd-master - Ranger Raz: Need to fix default truststore and keystore type

Replace hard coded jks with KeyStore.getDefaultType for initialising the default store type.

CDPD-60518: Introduce config within Ranger to control retention period of x_trx_log data

Add config within Ranger to control retention period of x_trx_log table data.

CDPD-60268: [7.2.18 - CLONE] - RangerJSONAuditWriter creates new log file for writing ranger audits as JSON every time there is an Exception

Fixes unnecessary new audit log files from getting created.

CDPD-59587: CLONE [7.2.18] - Ranger RMS for Ozone

Ranger RMS will support authorization for Ozone storage locations. RMS for Ozone will co-exist with Hive-HDFS ACL sync and provide authorization for both HDFS and Ozone file systems.

CDPD-59133: CLONE - Ranger[7.2.18] : Upgrade commons-configuration2 to 2.9.0 due to CVEs

As part of this fix, upgraded commons-configuration2 to 2.9.0

CDPD-58569: Ranger - Upgrade Guava to 32.0.1 due to CVE-2023-2976

Upgrade Guava library version to 32.0.1.

CDPD-58493: Ranger - Upgrade Netty Project to 4.1.94.Final due CVE-2023-34462

Upgrade Netty Project to 4.1.94.Final.

CDPD-57453: Atlas Error while writing audits to GCP Datalake

Removed the lib which was causing the conflict.

CDPD-57318: Ranger - Upgrade jackson-dataformat-xml to 2.13.5 due to multiple CVEs in woodstox

Use woodstox-core to 5.4.0 version.

CDPD-57018: Ranger - Upgrade aws-java-sdk to 1.12.367+

Upgrade aws-java-sdk to 1.12.481.

CDPD-56737: Ranger - Upgrade Tomcat to 8.5.89 due to CVE-2023-28709

Upgrade Tomcat to 8.5.89.

CDPD-56384: Ranger - Upgrade Spring LDAP to 2.4.1 due to high CVEs

Upgrade Spring LDAP to 2.4.1.

CDPD-56383: Ranger - Upgrade BeanShell to 2.1b5 due to high CVEs

Upgrade BeanShell to 2.1b5 by upgrading testNG to 7.0.0.

CDPD-56381: Ranger - Upgrade Apache Derby due to critical CVEs

Upgrade Apache Derby to 10.14.2.0.

CDPD-56343: Feature request for Ranger : More than 25 policies per page

This issue is fixed in ranger admin react UI.

CDPD-56300: Introduce config within Ranger to control retention period of x_auth_session data

Add config within Ranger to control retention period of x_auth_session table data.

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-55572: shell script to export, transform, import of ranger Roles for ranger replication

Shell script to export, transform, import of ranger Roles for ranger replication.

CDPD-55561: Ranger - Upgrade bcpkix-jdk15on to 1.70+ due to CVE-2019-17359

Upgrade bcpkix-jdk15on to 1.70.

CDPD-55459: Ranger - Upgrade Spring Framework to 5.3.27/6.0.8 due to CVE-2023-20863

Upgrade Spring Framework to 5.3.27.

CDPD-55419: Ranger - Upgrade json-smart to 2.4.10 due to CVE-2023-1370

Upgrade json-smart to 2.4.10.

CDPD-55050: Support SELF_OR_PREFIX resource matching scope in Ranger Authorization

API to find whether a user/group/role is authorized to the given operation on any resource of given type.

CDPD-53651: [UMBRELLA] Ranger Replication

Ranger Policy Replication support in Ranger.

CDPD-50564: Add/ Update Additional metric details for Ranger RMS

Add Additional Metrics for Ranger RMS.

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

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

CDPD-39939: [PAAS] Ranger RMS improvements

Added support for RMS in public cloud (AWS) to track s3 locations of Hive tables and databases.

CDPD-6087: RangerAuthorizationCoprocesor Unable to get remote Address

Issue already fixed in <https://issues.apache.org/jira/projects/RANGER/issues/RANGER-3758>

<https://jira.cloudera.com/browse/CDPD-45528> Log level changed from info to trace.

Apache Patch Information

- RANGER-4655
- RANGER-4611
- RANGER-4461
- RANGER-4407
- RANGER-4353
- RANGER-4342
- RANGER-4308
- RANGER-4262
- RANGER-4257
- RANGER-4255
- RANGER-4245
- RANGER-4242
- RANGER-4241
- RANGER-4220
- RANGER-4212
- RANGER-4165
- RANGER-4025
- RANGER-3758

Fixed Issues in Schema Registry

Review the list of Schema Registry issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-56890: New schemas cannot be created following an upgrade

Schemas can be created again after an upgrade even if the latest version of the schema was deleted before the upgrade.

CDPD-58265: Schema Registry Client incorrectly applies SSL configuration

The Cloudera distributed Schema Registry Java client applies the SSL configurations correctly even with concurrent access in Jersey clients.

CDPD-49470: Schema Registry Client retries requests more than the configured maxAttempts when multiple URLs are used

The Cloudera distributed Schema Registry Java client handles each request as one attempt, and does not attempt more retries based on the number of Schema Registry server URLs anymore.

OPSAPS-68139: Schema Registry does not apply cluster wide Kerberos principal mapping by default

The Schema Registry Kerberos Name Rules property is now empty by default. Schema Registry now automatically applies the cluster-wide auth-to-local (ATL) rules by default. During an upgrade, the previously configured value is preserved. If you have been using the default or a custom value, you must manually clear the property following an upgrade to transition to the new default value.

CDPD-55381: Schema Registry issues authentication cookie for the authorized user, not for the authenticated one

Schema Registry authentication cookie contains the correct authenticated user, even if the authenticated and the authorized users are different. Authenticated and authorized users can be different in scenarios where Schema Registry is used behind Knox.

CDPD-60160: Schema Registry Atlas integration does not work with Oracle databases

The Schema Registry Atlas integration works correctly when Oracle is used as the database of Schema Registry.

CDPD-59015: Schema Registry does not create new versions of schemas even if the schema is changed

The new fingerprinting version introduced in Cloudera Runtime 7.2.18 solves this issue.

CDPD-58990: The `getSortedSchemaVersions` method should order by version number and not by `schemaVersionId`

Schemas are ordered correctly by their version number during validation instead of their ID number.

CDPD-58949: Schemas are deduplicated during import

Importing works correctly and Schema Registry does not deduplicate imported schema versions.

Fixed Issues in Apache Solr

There are no fixed issues for Solr in Cloudera Runtime 7.2.18.

Apache Patch Information

None

Fixed Issues in Spark

Review the list of Spark issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-3038: Launching pyspark displays several HiveConf warning messages

When pyspark starts, several Hive configuration warning messages are displayed, similar to the following:

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

These errors can be safely ignored and this issue has been fixed.

CDPD-65717: SPARK-46793 Revert S3A endpoint fixup logic of SPARK-35878

SPARK-46793. Revert S3A endpoint fixup logic of SPARK-35878

CDPD-64638: Slowness / broadcast timeout issues due to SPARK-33290: REFRESH TABLE should invalidate cache even though the table itself may not be cached (Spark 2.4.8)

Slowness / broadcast timeout issues could occur due to SPARK-33290 in case of Spark 2.4.8. A new legacy `spark.sql.legacy.refreshOnlyCachedTables` feature flag has been introduced to restore the behavior prior to Spark 2.4.8. If `spark.sql.legacy.refreshOnlyCachedTables` is set to false (default), REFRESH TABLE should invalidate cache even though the table itself may not be cached, this was introduced with SPARK-33290 in Spark 2.4.8. When set to true, restore the behavior prior to Spark 2.4.8. I have manually tested with customer data which caused timeout / slowness issues.

CDPD-64546: Performance: Spark TPCDS Queries are slower in 7.2.18 compared to 7.2.17

Fixed with disabling checksum on the client side while reading data. The read performance is similar as earlier showing no regressions.

CDPD-61564: Spark - Caused by: `java.lang.NoClassDefFoundError: org/datanucleus/store/query/cache/QueryCompilationCache`

Upgraded datanucleus-core dependency to 5.2.10

CDPD-57535: Revert: CDPD-48171: Temporary workaround pinning snakeyaml to 2.0 not vulnerable to CVE-2022-1471

Reverted back from snakeyaml 2.0. The snakeyaml's Representer constructor has been added back. The other reverted constructors can be found here: <https://bitbucket.org/snakeyaml/snakeyaml/commits/3e755d254aeaa902675053047fd53368a175565a/raw>

CDPD-58558: Simple DML insert into table via spark3-shell sparks.sql is creating orphan spark_process in atlas

Does not create spark_process entity in case of INSERT INTO ... VALUES ... Only the INSERT INTO ... SELECT ... action may create spark_process entity in Atlas based on these official documentations: <https://docs.cloudera.com/cdp-private-cloud-base/7.1.8/atlas-reference/topics/atlas-spark-actions.html> <https://docs.cloudera.com/runtime/7.2.17/atlas-reference/topics/atlas-spark-actions.html>

CDPD-58191: Spark - Upgrade kubernetes library to 5.7.4/5.8.1/5.10.2/5.11.2+ due to CVE-2021-4178

Upgraded kubernetes-client dependency to 5.7.4

CDPD-58080: Backport SPARK-32951 to Spark 2

SPARK-32951 Foldable propagation from Aggregate

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

Added CREATE VIEW lineage support to Spark Atlas Connector for Spark3 which is required for Iceberg tables

CDPD-56342: Upgrade Parquet to 1.12.3 in Spark

Upgraded Parquet dependency to 1.12.3

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-55116: Fix Spark vulnerability CVE-2023-22946

This fix is blacklisting “spark.submit.deployMode” and “spark.submit.proxyUser.allowCustomClasspathInClusterMode” spark configurations in Livy create session REST API. We have added a new Livy configuration “livy.server.session.allow-custom-classpath” to allow custom class path. In order to disable or rollback this fix, we can add “livy.server.session.allow-custom-classpath” as “true” in Livy configuration via the CM safety valve.

CDPD-44454: MAPREDUCE-7432. Make manifest committer default on abfs and gcs stores

MAPREDUCE-7432. Make manifest committer default on abfs and gcs stores

CDPD-44227: Ranger improvement - Roles Import/export API for ranger admin

Add Roles Import/export API for ranger admin

Apache patch information

- SPARK-46793
- SPARK-39441
- SPARK-32951
- LIVY-975
- MAPREDUCE-7432

Fixed Issues in Spark3

Review the list of Spark3 issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-60190: Backport SPARK-39441

[SPARK-39441] Speed up DeduplicateRelations

CDPD-58191: Spark - Upgrade kubernetes library to 5.7.4/5.8.1/5.10.2/5.11.2+ due to CVE-2021-4178

Upgraded kubernetes-client dependency to 5.7.4

CDPD-57535: Revert: CDPD-48171: Temporary workaround pinning snakeyaml to 2.0 not vulnerable to CVE-2022-1471

Reverted back from snakeyaml 2.0. The snakeyaml's Representer constructor has been added back. The other reverted constructors can be found here: <https://bitbucket.org/snakeyaml/snakeyaml/commits/3e755d254aeaa902675053047fd53368a175565a/raw>

CDPD-56342: Upgrade Parquet to 1.12.3 in Spark

Upgraded Parquet dependency to 1.12.3

CDPD-55116: Fix Spark vulnerability CVE-2023-22946

This fix is blacklisting “spark.submit.deployMode” and “spark.submit.proxyUser.allowCustomClasspathInClusterMode” spark configurations in Livy create session REST API. We have added a new Livy configuration “livy.server.session.allow-custom-classpath” to allow custom class path. In order to disable or rollback this fix, we can add “livy.server.session.allow-custom-classpath” as “true” in Livy configuration via the CM safety valve.

Apache patch information

None

Fixed Issues in Apache Sqoop

Review the list of Sqoop issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-44397: Implement ORC support in Sqoop-Connector-Teradata component

A new version of Cloudera Connector Powered by Teradata version 1.8.5.1c7 is released which includes ORC support in the Sqoop-Connector-Teradata component. You can use Teradata Manager to import data from the Teradata server to Hive in ORC format.

CDPD-47175: Sqoop Hive import with ORC file fails with ClassCastException

The import process of Sqoop to ORC file has been updated. Whenever an unsupported conversion is attempted, Sqoop now provides a comprehensive error message describing the issue.

Sqoop can now import the following data types:

- Byte, Short, Int, Long, Float, Double from the same RDBMS types
- BigDecimal to Long, Double, String
- Date, Timestamp to String, Date, Timestamp

CDPD-56523: Sqoop does not take --hive-compute-stats option into account for hs2-url Hive imports

Sqoop now considers the --hive-compute-stats option for Hive imports when hs2-url parameter is used.

CDPD-58538: Oozie should upload and use the config files from sqoop-conf/managers.d when available

Previously, Oozie did not honor Sqoop's managers.d configurations and extra connector JARs from the lib folder, but now both are automatically available in Oozie's Sqoop action, allowing users to seamlessly utilize connectors like the Sqoop Teradata connector without the need for manual configuration updates or copying JARs to the Workflow's lib folder.

CDPD-59557: Secure options to provide the Hive password for Sqoop Hive imports

This fix introduces secure options that you can use to provide the Hive password during Sqoop-Hive imports instead of the earlier way of providing the password as plaintext in the command-line interface.

CDPD-59710: Fix time stamp conversion issue when exporting Parquet

When available, Sqoop will incorporate the writer's time zone metadata from the Parquet file during the export operation.

CDPD-61547: Sqoop should not close 'System.out' and 'System.err'

In certain cases the Sqoop process closed the 'sysout' and 'syserr' streams making it impossible to write to these if Sqoop manually used in a custom JVM.

CDPD-63723: Sqoop should determine files as Parquet by PAR1 in header

Sqoop now looks at the first 4 bytes of a file instead of 3 bytes to determine if the file is a Parquet file or not

CDPD-63915: Sqoop Teradata export fails if the source table is empty

Fixed the issue where Sqoop Teradata export failed if the source table was empty

Apache patch information

None

Fixed Issues in Streams Messaging Manager

Review the list of Streams Messaging Manager issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-53437: Alert notification messages do not contain an alert cause

In some cases, the resources and the attributes of the resource that triggered the alert were not included in the alert message. Typically, this happened when the alert policy defined a WITH ANY clause. From now on, resources that trigger the alert are always included in the alert message.

OPSAPS-69481: Anchor links are broken on Knox enabled clusters

Links that appear on the **Overview** page when hovering over producers, consumers, or partitions are no longer broken on Knox-enabled clusters.

Fixed Issues in Streams Replication Manager

Review the list of Streams Replication Manager issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-60426: Configuration changes are lost following a rolling restart of the service

SRM no longer fails to apply configuration changes if it is restarted with a rolling restart.

OPSAPS-67772: SRM Service metrics processing fails when the noexec option is enabled for /tmp

The SRM Service Kafka Streams application now uses the Kafka Streams state directory to extract the RocksDB .so files.

OPSAPS-67738: SRM Service role's Remote Querying feature does not work when the noexec option is enabled for /tmp

The SRM service does not add Netty native libraries to /tmp by default as streams.replication.manager.service.netty.native.working.dir configuration was introduced.

OPSAPS-67742: The SRM Service role fails to start if properties are added to Additional Configs For Streams Application Running Inside SRM Service

The SRM Service role no longer fails to start if properties are added to the Additional Configs For Streams Application Running Inside SRM Service configuration. It is also possible to configure the internal Kafka Streams application of the SRM Service role.

Apache patch information

None

Fixed Issues in Apache Tez

Review the list of Tez issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-60837: Tez - Upgrade Guava to 32.0.1 due to CVE-2023-2976

Upgrade Guava to 32.0.1 due to CVE-2023-2976

CDPD-53825: Tez - Upgrade jettison to 1.5.4 due to CVE-2023-1436

Upgraded jettison to 1.5.4

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 7.2.18.

COMPX-15361: Backport YARN-11590 to cdpd-master: RM process stuck after calling confStore.format() when ZK SSL/TLS is enabled, as netty thread waits indefinitely

Fixed the issue when SSL/TLS was enabled for YARN with Curator leader-elect, a Netty thread rendered RM unresponsive.

COMPX-15323: YARN Client unit tests are failing because of "NoClassDefFoundError: org/junit/jupiter/api/TestInfo"

Unit test dependency added to the pom.

COMPX-15308: YARN-11578 Fix performance issue of permission check in verifyAndCreateRemoteLogDir

Performance issue fix in LogAggregation's verifyAndCreateRemoteLogDir.

COMPX-15299: YARN-9954 Configurable max application tags and max tag length

Configurable max application tags and length.

COMPX-15289: Make t9000 compatible with mixed resource mode allocation

Now the fine plugin is disabled in mixed mode

COMPX-15288: [7.2.18] QM Database migration breaks for clusters where H2 database prior to migration had duplicate config_sets for a namespace, version pair

Fix tracked under COMPX-15216

COMPX-15272: Yarn - Queue Manager cannot load correctly on 7.2.18

Changes done in <https://jira.cloudera.com/browse/COMPX-15432> to make embedded H2 db as the default db again removing the requirement on a postgres db for QM in 7.2.18

COMPX-15111: Backport HADOOP-18870: CURATOR-599 change broke functionality introduced in HADOOP-18139 and HADOOP-18709

Hadoop Common can be configured now to communicate via SSL/TLS with ZK.

COMPX-14929: Backport YARN-11468: Zookeeper SSL/TLS support

Zookeeper SSL/TLS support enabled for YARN.

COMPX-14928: Backport HADOOP-18709: Add curator-based ZooKeeper communication support over SSL/TLS into the common library

Hadoop Common can be configured now to communicate via SSL/TLS with ZK.

COMPX-14855: Backporting YARN-11535 (Remove jackson-dataformat-yaml dependency)

Removed Jackson-dataformat-yaml to fix CVEs

COMPX-14759: Queue Manager - Upgrade commons-configuration2 to 2.9.0 due to CVEs

Updated commons-configuration2 to 2.9.0 for 7.1.9 CHF2

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

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

COMPX-14064: Default ULF for dynamic queue template should be -1 in Weight mode

Set Default ULF for dynamic queue template to be -1

COMPX-13056: QM - Upgrade Grizzly-http to 4.0.0 due to CVE-2017-1000028

Removed grizzly dependency

COMPX-9022: YARN Scaling metrics API should return queue statistics

The API was extended with new debug infos

CDPD-64617: [7.2.18][terminate-flow] yarn cmd -getServiceState trying to connect to an intentionally down host in ranger-post

Due to multi-comp execution on the same cluster some services are in stopped state. When recreated the scenario with one master node down and all services running this issue is gone and tests started executing.

CDPD-57948: [7.1.9 ZDU Simulation] Hive Query is failing when YARN is into rolling restart

YARN-side fix is implemented and backported to cdpd-master and 7.1.9.x

Apache patch information

- MAPREDUCE-7468
- MAPREDUCE-7446
- MAPREDUCE-7432
- YARN-11578
- YARN-9954
- YARN-11520
- YARN-11551
- YARN-11535
- YARN-11545
- YARN-11533
- YARN-11490
- YARN-11621

Fixed Issues in Zeppelin

Review the list of Zeppelin issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-63306: Upgrade netty to 4.1.100.Final due to CVE-2023-44487

Netty is upgraded to 4.1.100.Final

Apache patch information

None

Fixed Issues in Apache ZooKeeper

Review the list of ZooKeeper issues that are resolved in Cloudera Runtime 7.2.18.

CDPD-62448: Explicit handling of DIGEST-MD5 vs GSSAPI in quorum auth

Explicit handling of DIGEST-MD5 vs GSSAPI in quorum auth has been added. Fixing CVE-2023-44981 with the backport of ZOOKEEPER-4753.

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

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

CDPD-51890 Replace logback with reload4j

This issue is fixed now.

OPSAPS-68466 Enable ZooKeeper Client port unification in CM

Added a new ZooKeeper param to CM which allows the user to enable TLS connections on the unsecure client port. Param name: client.portUnification.

Apache Patch Information

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

Known Issues In Cloudera Runtime 7.2.18

You must be aware of the known issues and limitations, the areas of impact, and workaround in Cloudera Runtime 7.2.18.

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-56085: [Impala Iceberg] LOAD DATA INPATH to Iceberg_table creates a temporary hive_table with name <iceberg_table_name>_tmp* and then marks it as DELETED in Atlas

Running a query like "LOAD DATA INPATH to iceberg_table", creates a temporary hive_table with name <iceberg_table_name>_tmp* and then marks it as DELETED in Atlas. So in Atlas, a deleted entity is created corresponding to the temporary table "<iceberg_table_name>_tmp*".

Tag added to the File system (HDFS) entity will not be propagated to the Iceberg table, user has to manually add to the iceberg_table, since the tag propagation is broken due to the deleted table in the flow.

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

CDPD-67112: Import transforms do not work as expected when replacing a string which already has ":"

None

CDPD-67022: Export/Import: When a user with export/import permissions does not have the permission to create/read/write/update entity, the import operation fails with 403 error

The entity permission for __AtlasAuditEntry has to be present.

CDPD-67020: When a user has export/import permissions but no other permission on entity (read/write/update), export operation throws an error, import operation fails

The entity permission for AtlasServer and __ExportImportAuditEntry has to be present.

CDPD-65806: After upgrading from Cloudera Runtime 7.2.17 to 7.2.18, not all Iceberg table relationships are visible in the entity details page

None

OPSAPS-68461: Update GC and JVM options for Atlas service for supporting JDK17 in main Atlas CSD

Existing ATLAS_OPTS does not work for JDK17. You must manually update ATLAS_OPTS.

DOCS-19084: 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).
- During the Zookeeper service upgrade process, any API requests routed to the Passive node shall result in 503 error, until the Active instance is up.
- During the Rolling upgrade, when the Hbase backup is being carried out, Atlas service is stopped and there can be a minimal disruption till the HBase backup process is completed. Atlas service gets restarted once the HBase backup process is completed.

CDPD-62973: Change in audits behavior in Cloudera Runtime 7.2.18 deployment.

When the value of differential audits is set as true, the audit information is not segregated based on the user which is firing the query. The HMS service user information includes details of the service user. When differential audit is enabled, only the difference between the two subsequent audits is logged, but in this case, there is no change in the data which is retrieved from HS2 and HMS, which does not create the audit. The user information is audited fine when differential audit is disabled

DOCS-19610: After upgrading from Cloudera Runtime 7.2.17 to 7.2.18, not all Iceberg table relationships are visible in the entity details page

The following entities are affected:

- relationshipAttributes
- hive partition_spec
- database details missing
- ddl_queries

CDPD-63397: During Data Lake upgrade, Atlas authorization is denied

When rolling upgrade is performed, there might be a scenario where Ranger Admin could be undergoing upgrade by itself and hence the policy download could be affected.

During this period, access might be denied for certain Atlas entities. This issue is resolved once Ranger Admin is up and the policies are downloaded.

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

None

CDPD-7781: TLS certificates not validated on Firefox

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

None

CDPD-6675: Irregular qualifiedName format for Azure storage

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

None

CDPD-5933 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 "@" 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

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

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

There are no known issues for Cloud Connectors in Cloudera Runtime 7.2.18.

Known issues in Cruise Control

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

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.

Known Issues in Apache HBase

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

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

Stop HBase using Cloudera Manager before deleting an 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.

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.

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

Snappy compression with `/tmp` directory mounted with `noexec` option

Using the HBase client applications such as `hbase hfile` on the cluster with Snappy compression could result in `UnsatisfiedLinkError`.

Add `-Dorg.xerial.snappy.tmpdir=/var/hbase/snappy-tempdir` to Client Java Configuration Options in Cloudera Manager that points to a directory where `exec` option is allowed.

Known Issues in HDFS

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

CDPSDX-5302: Avoiding long delay on the HBase master does not happen during upgrade.

1. Log in to Cloudera Manager
2. Select the HDFS service
3. Select Configurations tab
4. Search for `hdfs-site.xml`.
5. Set `ipc.client.connect.timeout = 5000`
6. Set `ipc.client.connect.max.retries.on.timeouts = 5`
7. Click Save

The above configuration changes ensures that:

1. The long delay on the HBase master does not happen during upgrade.
2. The long delay on the HBase master recovery does not happen during upgrade.

CDPD-65530: HDFS requests throw `UnknownHostException` during OS upgrade

During the VM replacement as part of OS upgrade, every new node gets a new IP Address, and if the old IP address is cached somewhere, HDFS requests fail with `UnknownHostException` and it recovers after sometime (10 mins max).

The issue is seen during COD and DL ZDU.

None.

CDPD-67230: Rolling restart can cause failed writes on small clusters

In a rolling restart, if the cluster has less than 10 datanodes existing writers can fail with an error indicating a new block cannot be allocated and all nodes are excluded. This is because the client has attempted to use all the datanodes in the cluster, and failed to write to each of them as they were restarted. This will only happen on small clusters of less than 10 datanodes, as larger clusters have more spare node to allow the write to continue.

None.

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](#))

Known Issues in Apache Hive

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

CDPD-60418: Hive beeline queries are failing with a `org.apache.thrift.transport.TTransportException` error

When running Hive beeline queries on a custom Data Engineering Data Hub cluster on a Google Cloud Platform (GCP) environment with Ranger Authorization Service (RAZ), the queries failed. It was noticed that the cluster has multiple connection failures from HiveServer (HS2) to dependent services.

On further analysis, it was observed that the issue is related to the `fs.s3a.ssl.channel.mode=open` `ssl` parameter that was provided in the custom cluster template configuration. OpenSSL is not supported for RAZ on a GCP environment.

1. In Cloudera Manager of the custom DE Data Hub cluster, go to Clusters Hive Metastore Configuration
2. Search for Hive Service Advanced Configuration Snippet (Safety Valve) for `hive-site.xml` and remove the `fs.s3a.ssl.channel.mode` parameter.
3. Restart the Hive Metastore service.

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 7.2.18

CDPD-64541, CDPD-63617: Creating managed tables using Hue Importer fails on RAZ-enabled GCP environments

On Google Cloud Platform (GCP) environments, creating managed tables in both Hive and Impala dialects fails and temporary (tmp) tables are dumped (created). This is most likely because Hive and Impala cannot load data inpath from Google Storage (outside of Hue).

None.

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***] / [***DIRECTORY-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-DIRECTORY***] [***COMPLETE-PATH-TO-S3-BUCKET***] / [***TARGET-DIRECTORY***]
```

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.

Known issues in 7.2.16**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**

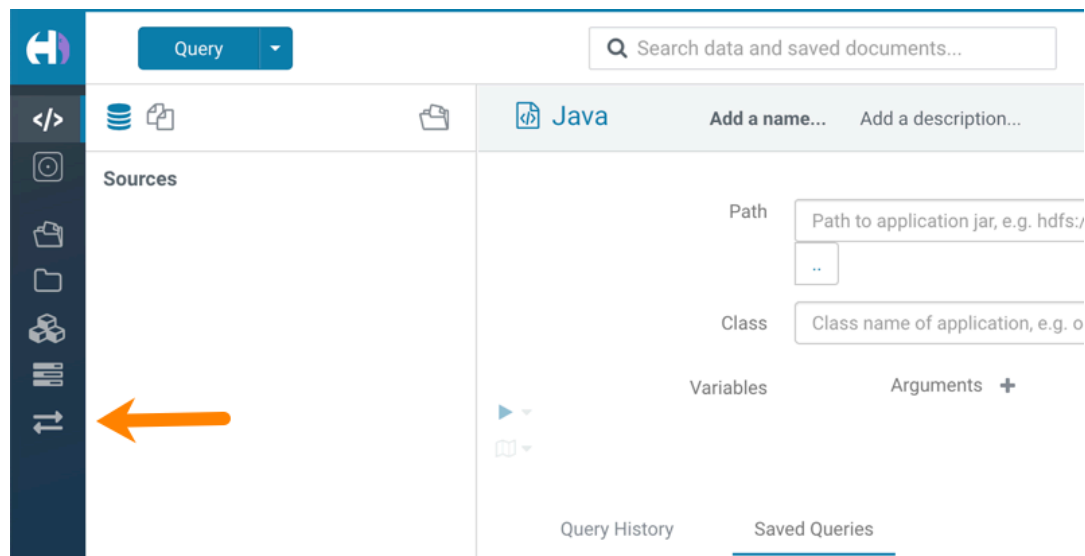
When you run Impala queries in a batch mode, you encounter failures with a Kerberos error even if the keytab is configured correctly. This is because submitting Impala, Sqoop, Pig, or pyspark queries in a batch mode launches a shell script Oozie job from Hue and this is not supported on a secure cluster.

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>;
```

CDPD-66305: Do not turn on the optimized Iceberg V2 operator in 7.2.18.0

In this release, the optimized Iceberg V2 operator is disabled by default due to a correctness issue. The correct setting for the property that turns off the operator is `DISABLE_OPTIMIZED_ICEBERG_V2_READ=true`.

Accept the default setting of the V2 operator. Do not change the setting from true to false.

CDPD-64629: Performance degradation of Iceberg tables compared to Hive tables

Cloudera testing of Iceberg and Hive tables using the Hive TPC-DS 1 Tb dataset (Parquet) revealed a slower performance executing a few of the queries in TPCDS. Overall performance of Iceberg executing queries on Hive external tables of Iceberg is faster than Hive.

Known Issues in Apache Impala

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

CDPD-57989: MERGE INTO Query fails on tables with non-nullable columns.

None

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/r/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 HDF
S file hdfs://nameservice1/warehouse/tablespace/managed/hive/<da
tabase>/<table>/xxxxx
Error(2): No such file or directory Root cause: RemoteException:
File does not exist: /warehouse/tablespace/managed/hive/<data
base>/<table>/xxxx
```

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

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](#)

Severity

7.5 (High) [CVSS:3.1/AV:N/AC:H/PR:L/UI:N/S:U/C:H/I:H/A:H](#)

Releases affected

- CDP Private Cloud Base 7.0.3, 7.1.1, 7.1.2, 7.1.3, 7.1.4, 7.1.5 and 7.1.6
- CDP Public Cloud 7.0.0, 7.0.1, 7.0.2, 7.1.0, 7.2.0, 7.2.1, 7.2.2, 7.2.6, 7.2.7, and 7.2.8
- All CDH 6.3.4 and lower releases

Impact

Unauthorized access

Users affected

Impala users of the affected releases

Action required

Upgrade to a CDP Private Cloud Base or CDP Public Cloud version containing the fix.

Addressed in patch/release/hotfix

- CDP Private Cloud Base 7.1.7
- CDP Public Cloud 7.2.9 or higher versions

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 alltypeptyny a1
  INNER JOIN alltypesagg a2 ON a1.smallint_col = a2.year AND fals
e
  RIGHT JOIN alltypes a3 ON a1.year = a1.bigint_col;
+---+
| Explain String |
+---+
| Estimated Per-Host Requirements: Memory=1.00KB VCores=1 |
| 00:EMPTYSET |
+---+
```

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

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

Add `-minidump_size_limit_hint_kb=size` to set a soft upper limit on the size of each minidump file. If the minidump file would exceed that limit, Impala reduces the amount of information for each thread from 8 KB to 2 KB. (Full thread information is captured for the first 20 threads, then 2 KB per thread after that.) The minidump file can still grow larger than the "hinted" size. For example, if you have 10,000 threads, the minidump file can be more than 20 MB.

IMPALA-4978: Impala requires FQDN from hostname command on Kerberized clusters

The method Impala uses to retrieve the host name while constructing the Kerberos principal is the `gethostname()` system call. This function might not always return the fully qualified domain name, depending on the network configuration. If the daemons cannot determine the FQDN, Impala does not start on a Kerberized cluster.

Test if a host is affected by checking whether the output of the `hostname` command includes the FQDN. On hosts where `hostname` only returns the short name, pass the command-line flag `##hostname=fully_qualified_domain_name` in the startup options of all Impala-related daemons.

IMPALA-6671: Metadata operations block read-only operations on unrelated tables

Metadata operations that change the state of a table, like `COMPUTE STATS` or `ALTER RECOVER PARTITIONS`, may delay metadata propagation of unrelated unloaded tables triggered by statements like `DESCRIBE` or `SELECT` queries.

Workaround: None

IMPALA-7072: Impala does not support Heimdal Kerberos

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

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

For example, when a user ingests data from a file into a partition of an existing table using Spark, if `spark.hadoop.hive.stats.autogather` is not set to `false` explicitly, `numRows` associated with this partition would be 0 even though there is at least one row in the file. To avoid this, the workaround

is to set "spark.hadoop.hive.stats.autogather=false" in the "Spark Client Advanced Configuration Snippet (Safety Valve) for spark-conf/spark-defaults.conf" in Spark's CM Configuration section.

Technical Service Bulletins

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

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

Upstream JIRA

[IMPALA-7561](#)

Impact

Potential incorrect query results, due to incomplete dataset.

Action required

- **Upgrade (recommended)**

This is fixed in the newest versions of the Impala JDBC driver and the Impala ODBC driver, available at the following locations:

- Impala ODBC 2.6.12 - <https://www.cloudera.com/downloads/connectors/impala/odbc/2-6-12.html>
- Impala JDBC 2.6.20 - <https://www.cloudera.com/downloads/connectors/impala/jdbc/2-6-20.html>
- **Workaround**

Set the property "`FETCH_ROWS_TIMEOUT_MS`" to 0 if you are unable to use one of the newer versions of the respective drivers listed above. This way, the client can fetch the complete set of data without any issues. Setting the timeout to 0 effectively turns the fetch call into a blocking request which will not timeout and will wait till all the results are fetched. It will wait until all the results come through and/or the network layer timeouts.

This can be set at the Impala server level (via [Impala Daemon Query Options safety valve](#)), or in a pool used with Admission Control, or at the session level, or at the query level.

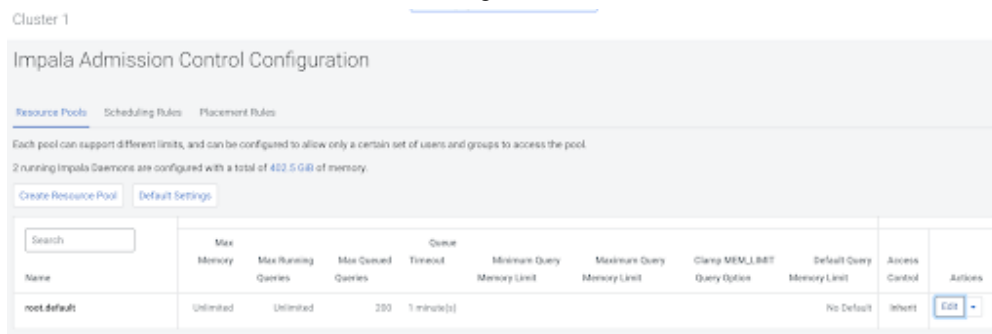
- On the command line, it can be changed at the server level as follows:

```
bin/start-impala-cluster.py --impalad_args="--default_query_options=fetch_rows_timeout_ms=0"
```

- To set the query option at the user session level: `SET fetch_rows_timeout_ms=0;`

- To set it in admission controls pools via CM:
 - Click Clusters Impala Admission Control Configuration .

You will see a screen similar to the following:



- Click Edit and scroll down to Default Query Options.
 - Click + and add the query option `FETCH_ROWS_TIMEOUT_MS = 0`
- In CDW, users can configure their Impala VW by the following steps:
 - Click the three dots and choose the Edit option.
 - Select the Impala coordinator tab within the Configurations tab.
 - Select flagfile in the drop-down menu.
 - Find the `default_query_options` key and add the following to the end of the value string: `FETCH_ROWS_TIMEOUT_MS = 0`. Make sure to add a comma before adding this to the value string.

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](#)

Impact

Incorrect results returned from certain Impala queries.

Action required

- Preferred Solution/Upgrade**

Please contact Cloudera Support for raising a Hotfix request until a release with the fix is available.

- Workaround**

None

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.

CDPD-64652: During CDH + OS rolling upgrade knox admin api access fails with 403 ACL authorization failures

During OS upgrades, attempts to access Knox on the host being upgraded may produce occasional 403 HTTP responses.

Since the cause is the unavailability of underlying OS service(s), wait and retry the failed request(s).

CDPD-60630: Knox redirecting Yarn Node Manager URLs to http instead of https

While viewing the yarn application logs on YARN RM UI via Knox, we can see that Knox is redirecting the NM URL to HTTP instead of HTTPS, as YARN is running on TLS/SSL.

```
https://<knox-gateway>/gateway/cdp-proxy/yarn/nodemanager/node?scheme=http&host=some.url&port=8044
```

Change *scheme=https* in the URL, and the page loads without issues.

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 7.2.18.

Known Issues in Apache Ranger

There are no known issues for Ranger in Cloudera Runtime 7.2.18.

Known Issues in Schema Registry

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

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 s://ranger-1.cloudera.com:6182/`

4. Restart the Schema Registry service.

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.ConnectionImplementat
ion.locateRegion(ConnectionImplementation.java:867)
        at org.apache.hadoop.hbase.client.ConnectionImplementati
on.relocateRegion(ConnectionImplementation.java:850)
        at org.apache.hadoop.hbase.client.ConnectionImplementat
ion.locateRegionInMeta(ConnectionImplementation.java:981)
        at org.apache.hadoop.hbase.client.ConnectionImplementa
tion.locateRegion(ConnectionImplementation.java:870)
        at org.apache.hadoop.hbase.client.RpcRetryingCallerWith
ReadReplicas.getRegionLocations(RpcRetryingCallerWithReadReplica
s.java:319)
        ... 21 more
Caused by: org.apache.hadoop.hbase.client.RetriesExhaustedExcept
ion: Failed contacting masters after 1 attempts.
Exceptions:
java.io.IOException: Call to address=quasar-bmyccr-2.quasar-bmy
ccr.root.hwx.site/172.27.19.4:22001 failed on local exception: j
ava.io.IOException: java.lang.RuntimeException: Found no valid a
uthentication 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.ZK
ConnectionRegistry'
```

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-fc9dacc265fabfc5
00b92112712505e3_pid{{PID}}.hprof ...
Unable to create /data/tmp/solr_solr-SOLR_SERVER-fc9dacc265fab
fc500b92112712505e3_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.getCor
eStatus(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 Authenti
cation exception: User: xyz@something.example.com is not allowed
to impersonate xyz@something.example.com
[c:RTOTagMetaOdd s:shard3 r:core_node11 x:RTOTagMetaOdd_shar
d3_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.

CrunchIndexerTool does not work out of the box if /tmp is mounted noexec mode

When you try to run CrunchIndexerTool with the /tmp directory mounted in noexec mode, It throws a snappy-related error.

Create a separate directory for snappy temp files which is mounted with EXEC privileges and set this directory as the value of the org.xerial.snappy.tmpdir java property as a driver java option.

For example:

```

export myDriverJarDir=/opt/cloudera/parcels/CDH//lib/solr/contri
b/crunch;export myDependencyJarDir=/opt/cloudera/parcels/CDH//
lib/search/lib/search-crunch;export myDriverJar=$(find $myDriv
erJarDir -maxdepth 1 -name 'search-crunch-*.jar' ! -name '*-job.
jar' ! -name '*-sources.jar');export myDependencyJarFiles=$(find
$myDependencyJarDir -name '*.jar' | sort | tr '\n' ',' | head
-c -1);export myDependencyJarPaths=$(find $myDependencyJarDir
-name '*.jar' | sort | tr '\n' ':' | head -c -1);export HADOOP_
CONF_DIR=;spark-submit --master local --deploy-mode client --
driver-library-path /opt/cloudera/parcels/CDH//lib/hadoop/lib/
native/ --jars $myDependencyJarFiles --driver-java-options ' -
Dorg.xerial.snappy.tmpdir=/home/systest/tmp ' --class org.apa
che.solr.crunch.CrunchIndexerTool $myDriverJar --input-file-form
at=avroParquet --input-file-reader-schema search-parquetfile/par
quet-schema.avsc --morphline-file /tmp/mrTestBase.conf --pipelin
e-type spark --chatty hdfs://[***HOSTNAME***]:8020/tmp/parquetfi
leparsertest-input

```

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.httpserver.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).

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.

Users with insufficient Solr permissions may encounter a blank 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 they receive a blank Admin UI with no information.

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:

- Panel with security info in admin UI's dashboard
- Incremental backup mode
- Schema Designer UI
- 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)

Limitations

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 them. Because of this, blockcache writing is disabled by default.

Default Solr core names cannot be changed

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

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. Because of this, you need to reindex collections that were created with Solr 6 or earlier.

Known Issues in Apache Spark

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

CDPD-64788: Inserting Bloom Filters in join operations for Spark 3.4

When `spark.sql.optimizer.runtime.bloomFilter.enabled` is enabled in Spark 3.4 (CDP 7.2.18), it causes considerable improvement in many queries but may cause regression in a few others. As the improvement is more significant, this behavior is retained in Spark 3.4 in Cloudera Spark versions.

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.

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. Stop using direct mode. Do not use the --direct option in Sqoop import or export commands.

CDPD-3089: Avro, S3, and HCat do not work together properly

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

Parquet columns inadvertently renamed

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

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

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

Importing multiple megabytes per row before initial-page-run check (ColumnWriter) can cause OOM. Also, rows that vary significantly by size so that the next-page-size check is based on small rows, and is set very high, followed by many large rows can also cause OOM.

None

Known issues in Streams Messaging Manager

Learn about the known issues 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

Importing large connector configurations/ flow.snapshots reduces the usability of the Streams
 Messaging Manager's Connectors page 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.

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

A fresh install of 7.2.18 of YARN Queue Manager does not allow user to bypass the Setup Database screen for YARN Queue Manager

YARN Queue Manager in Cloudera Data Platform (CDP) Private Cloud Base 7.2.18 does not require you to install a PostGres database, therefore users should not see the Setup Database screen and should be able to skip the Setup Database screen. With this known issue, users who are conducting a fresh install of 7.2.18 are not able to bypass the Setup Database screen as expected.

1. When conducting a fresh install of YARN Queue Manager in 7.2.18, you must ensure that you have both CDP and Cloudera Manager upgraded to 7.2.18.
2. When you reach the Setup Database screen in the Cloudera Manager installation wizard for Queue Manager, enter any dummy values for the following fields:
 - a. Database name: configstore
 - b. Database Username: dbuser
 - c. Database Password: dbpassword

YARN Queue Manager will not connect to PostGres with the above details and will fall back to the embedded database.

3. Run the following script command in a browser console to enable the Continue button:


```
document.querySelector('.btn.next').removeAttribute('disabled');
```
4. Click Continue and proceed with the YARN Queue Manager installation.
5. After installation is complete, SSH into the host that has Queue Manager installed, and run this command: `sed -i 's/migrationCompleted=true/migrationCompleted=false/' /var/lib/hadoop-yarn/migration.properties`



Note: Enable Queue Manager in the YARN configurations, and restart YARN.

6. Restart YARN Queue Manager.

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:cdw 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.

CDPD-67150: During the restore procedure it could happen that a node becomes unhealthy as the default YARN Capacity Scheduler configuration is loaded onto the node during the restart.

Perform an extra restart on the Resource Manager role that has the incorrect configuration to restore the correct configuration.

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

There are no known issues for Zookeeper in Cloudera Runtime 7.2.18.

None.

Behavioral Changes In Cloudera Runtime 7.2.18

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 7.2.18.

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 the way compaction initiator and cleaner threads are handled

Previous behavior:

The compaction initiator and cleaner threads are enabled and disabled by setting the `hive.compact.or.initiator.on` property to 'true' or 'false'.

New behavior:

A new property `hive.compactor.cleaner.on` is introduced that allows you to selectively enable or disable the cleaner thread.

This property is not listed and is set to 'false' by default. Add the property to Hive Metastore Server Advanced Configuration Snippet (Safety Valve) for `hive-site.xml` in Cloudera Manager to have the same out-of-the-box experience as in the previous version.

Also, ensure that you set the property to 'true' for the compactor to run on the HMS instance.

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 `javax.security.auth.useSubjectCredsOnly` JVM property is set to true by default. As a result, connectors are required to use their own credentials. This default change is true for newly provisioned clusters. On upgraded clusters, `useSubjectCredsOnly` remains set to false to ensure backwards compatibility.

Previous behavior:

The `javax.security.auth.useSubjectCredsOnly` JVM property was set to false

New behavior:

The `javax.security.auth.useSubjectCredsOnly` JVM property is set to true by default. As a result, connectors are required to use their own credentials. If you are migrating connectors from a cluster running a previous version of Runtime to a new cluster running 7.2.18 or later, you must ensure that credentials are added to the connector configuration when migrated. Otherwise, migrated connectors may not work on the new cluster.

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:

Added support for Ranger TagSync and UserSync HA

Previous behavior:

In versions < 7.2.18, port information for UserSync and Tagsync was not visible nor configurable.

New behavior:

In version 7.2.18 +, UserSync and TagSync port numbers are available at [Cloudera Manager Ranger Configuration](#) .

For example:

1. In Configuration Filters , highlight Ranger TagSync. Then, in Search, type http:

The following, configurable settings are now available:

Figure 2: Ranger TagSync port settings

CDEP Deployment from 2023-Aug-21 10:16

Cluster 1

RANGER-1

Actions ▾

Aug 21, 10:25 PM UTC

StatusInstancesConfigurationCommandsCharts LibraryAuditsRanger Admin Web UI 🔗Quick Links ▾

Filters (1)Role GroupsHistory & Rollback

Filters (1)

Clear All

> SCOPE

Clear

RANGER-1 (Service-Wide)

4

Ranger Admin

6

Ranger Tagsync

5

Ranger Usersync

7

> CATEGORY

Main

1

Advanced

0

Database

0

Logs

0

Monitoring

0

Performance

0

Ranger Tagsync TLS/SSL
Keystore File Alias

ranger.tagsync.service.https.attrib.
keystore.keyalias

ranger.tagsync.service.https.attrib.
keystore.keyalias

Ranger Tagsync Default Group

Tagsync HTTP Port

ranger.tagsync.service.http.port

ranger_tagsync_http_port

Ranger Tagsync Default Group

Tagsync HTTPS port

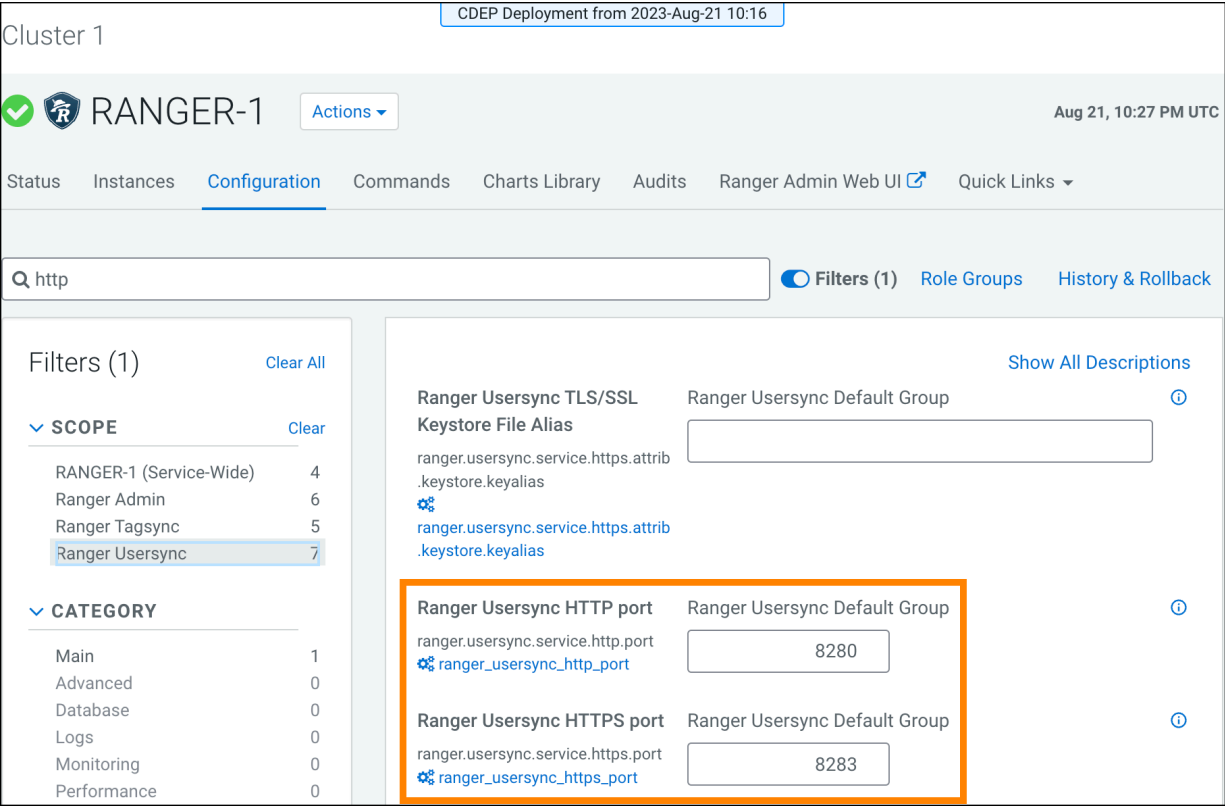
ranger.tagsync.service.https.port

ranger_tagsync_https_port

Ranger Tagsync Default Group

2. In Configuration Filters , highlight Ranger UserSync. Then, in Search, type http:
- The following, configurable settings are now available:

Figure 3: Ranger UserSync port settings



Deprecation Notices In Cloudera Runtime 7.2.18

Certain features and functionalities have been removed or deprecated in Cloudera Runtime 7.2.18. 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 Notice for DAS

Data Analytics Studio (DAS) has been deprecated and is no longer available in Cloudera Runtime starting with 7.2.18.

Removed component

Hue now replaces DAS. DAS features to support Hive and Tez such as running queries, defining HPL/SQL, the Job Browser, query explorer, query compare, and more, have been migrated to Hue, and the Hue Query Processor. After you upgrade to this release, you will not see the option to add the DAS service to your cluster. Cloudera recommends you use Hue for all use cases where you might have previously used DAS.

Deprecation Notices for Apache Kafka

Certain features and functionality in Apache Kafka are deprecated or removed in Cloudera Runtime 7.2.18. 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 Apache Oozie

Certain features and functionality in Apache Oozie are deprecated or removed in Cloudera Runtime 7.2.18. You must review these changes along with the information about the features in Oozie that will be removed or deprecated in a future release.

Deprecated

Oozie's Spark action

Due to the discontinuation and deprecation of Spark 2 in CDP 7.2.18, Cloudera decided to deprecate Oozie Spark actions, which are based on Spark 2. Consequently, Oozie's Spark actions will be disabled by default, and if you attempt to execute a Spark action, an error will be raised.

Starting from 7.2.18, Oozie introduces the new Spark 3 based Spark 3 actions.

Deprecation Notices for Spark 2

Spark 2 is deprecated in Cloudera Runtime 7.2.18. 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 7.2.18 is the last version in which Spark 2 will be supported, you will need to migrate to Spark 3.3.2 before you upgrade to a later version. See [Updating Spark 2 applications for Spark 3](#) linked below.

The following Spark 2 Data Hub templates were deleted:

- Data Engineering: Apache Spark, Apache Hive, Apache Oozie
- Data Engineering: HA: Apache Spark, Apache Hive, Apache Oozie
- Real-time Data Mart: Apache Impala, Hue, Apache Kudu, Apache Spark
- Data Discovery and Exploration

Related Information

[Using Spark 2 applications for Spark 3](#)

Deprecation Notices for Zeppelin

Starting from Cloudera Runtime 7.2.18, Zeppelin has been deprecated. While Zeppelin will continue to receive full support in 7.2.18, we recommend you to consider migrating to other Cloudera products. Detailed guidance on the recommended migration process will be provided shortly. Reach out to Cloudera Support or your dedicated Cloudera Account Team if you have inquiries or require further assistance.

Fixed Common Vulnerabilities and Exposures 7.2.18

Common Vulnerabilities and Exposures (CVE) that is fixed in this release.

- [CVE-2023-44487](#) - HTTP/2 Rapid Reset Vulnerability
- [CVE-2023-46120](#) - Netty
- [CVE-2023-36478](#) - Eclipse Jetty
- [CVE-2023-26048](#) - Eclipse Jetty
- [CVE-2023-26049](#) - Eclipse Jetty
- [CVE-2023-40167](#) - Eclipse Jetty
- [CVE-2023-36479](#) - Eclipse Jetty
- [CVE-2023-41900](#) - Eclipse Jetty
- [CVE-2023-38493](#) - Armeria
- [CVE-2021-31684](#) - Json-smart
- [CVE-2022-41881](#) - Netty
- [CVE-2023-34462](#) - Netty
- [CVE-2022-30187](#) - Azure-storage-blob, Azure-storage-queue
- [CVE-2023-24998](#) - Commons-fileupload

- [CVE-2019-10219](#) - Hibernate-validator
- [CVE-2020-10693](#) - Hibernate-validator
- [CVE-2019-16728](#) - DOMPurify
- [CVE-2019-25155](#) - DOMPurify
- [CVE-2023-37895](#) - Jackrabbit
- [CVE-2021-33813](#) - Jdom
- [CVE-2022-24823](#) - Netty
- [CVE-2023-31582](#) - Jose4j
- [CVE-2022-45688](#) - org.json
- [CVE-2023-5072](#) - org.json
- [CVE-2023-25613](#) - Apache Kerby
- [CVE-2020-28458](#) - Datatables
- [CVE-2019-20444](#) - Netty
- [CVE-2019-20445](#) - Netty
- [CVE-2019-16869](#) - Netty
- [CVE-2021-37136](#) - Netty
- [CVE-2021-37137](#) - Netty
- [CVE-2021-43797](#) - Netty
- [CVE-2021-21409](#) - Netty
- [CVE-2021-21290](#) - Netty
- [CVE-2023-34468](#) - Apache Nifi
- [CVE-2023-36542](#) - Apache Nifi
- [CVE-2023-40037](#) - Apache Nifi
- [CVE-2022-41917](#) - Opensearch
- [CVE-2023-34054](#) - Reactor-netty
- [CVE-2023-34062](#) - Reactor-netty
- [CVE-2022-31684](#) - Reactor-netty
- [CVE-2017-15288](#) - Scala-lang
- [CVE-2023-34478](#) - Apache Shiro
- [CVE-2023-20860](#) - Spring Framework
- [CVE-2023-20861](#) - Spring Framework
- [CVE-2023-20863](#) - Spring Framework
- [CVE-2023-34034](#) - Spring Security
- [CVE-2023-20862](#) - Spring Security
- [CVE-2023-32697](#) - Sqlite-jdbc
- [CVE-2023-41080](#) - Apache Tomcat
- [CVE-2023-42794](#) - Apache Tomcat
- [CVE-2023-42795](#) - Apache Tomcat
- [CVE-2023-45648](#) - Apache Tomcat
- [CVE-2023-4807](#) - Openssl
- [CVE-2023-2650](#) - Openssl
- [CVE-2023-3817](#) - Openssl
- [CVE-2023-5678](#) - Openssl
- [CVE-2022-45685](#) - Jettison
- [CVE-2022-45693](#) - Jettison
- [CVE-2023-1436](#) - Jettison
- [CVE-2021-21295](#) - Netty
- [CVE-2023-46750](#) - Apache Shiro
- [CVE-2022-46751](#) - Apache Ivy