Cloudera Runtime 7.3.1

Release Notes

Date published: 2020-07-28
Date modified: 2024-12-10

https://docs.cloudera.com/

Legal Notice

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.

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

This document provides you with the latest information about Cloudera Runtime 7.3.1. It includes improvements and describes new features, bug fixes, tech previews and more. For detailed information about the runtime components, see Cloudera documentation.

What's New In Cloudera Runtime 7.3.1

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

Spark 2 removed from Cloudera Runtime

Spark 3 is the default Spark version in Cloudera Runtime. Spark 2 has been removed and no longer available in 7.3.1.

Important:

Spark 3 contains a large number of changes from Spark 2.

Refer to *Upgrading Spark* for more information on upgrading Spark clusters to 7.3.1, and *Migrating Spark Applications* for more information on migrating your existing Spark applications between versions 2 and 3.

Upgrade CDP 7.2.17.200 through CDP 7.2.17.500 to CDP 7.3.1.

You can perform an in-place upgrade from 7.2.17.200 through CDP 7.2.17.500 to CDP 7.3.1. For more information, see the Upgrading documentation of CDP Public Cloud.

Upgrade CDP 7.2.18.0 through CDP 7.2.18.200 to CDP 7.3.1

You can perform an in-place upgrade from CDP 7.2.18.0 through CDP 7.2.18.200 to CDP 7.3.1. For more information, see the Upgrading documentation of CDP Public Cloud.

Rollback CDP 7.3.1 to a previous version

Rolling back to a previous version after upgrading to 7.3.1 is not supported. There is a recovery process for Data Lakes to return the cluster to its pre-upgrade state in case the upgrade fails. For more information, see Recovering from failed upgrades.

Related Information

Upgrading Spark

Migrating Spark Applications

What's New in Apache Atlas

Learn about the new features of Apache Atlas in Cloudera Runtime 7.3.1.

Cloudera Navigator Data Management upgrade to Apache Atlas

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

Business Metadata: Entity model extensions

This release of Atlas provides the ability for data stewards to add custom attributes to existing entity types and set their values on existing entities. This functionality allows an organization to extend its enterprise data model with curated master data attributes that have specific meaning for the business. Business Metadata attributes are defined centrally and can be used on designated entity types. Administrators can control who can view, add values to, and create or update set collections of Business Metadata attributes. Privileged users can add free-form values or select from predefined values to the attribute for a given entity. For more information, see Leveraging Business Metadata.

Bulk import of Business Metadata attribute associations

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

Bulk import of Glossary terms

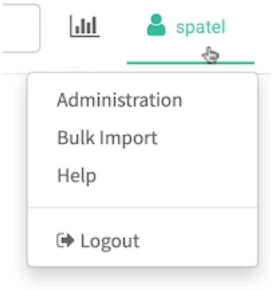
Atlas provides an interface to import a list of terms into existing Glossaries. The list can include any or all of the metadata associated with a given term. The list can be formatted as comma-separated values (.CSV) or Microsoft Excel (.XLS) formatted file. For more information, see Importing Glossary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

Administrator features have a home in the Atlas UI

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

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

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



Purge of deleted entities

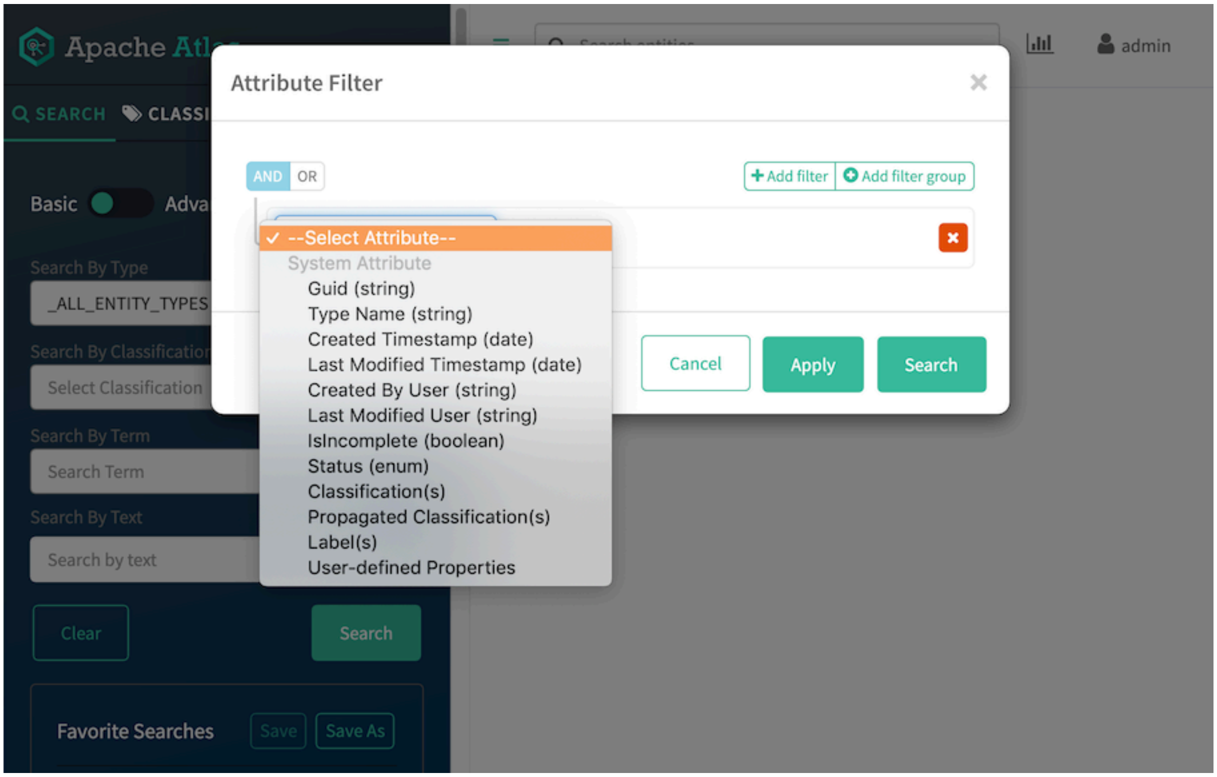
Atlas now provides the ability to clear the metadata for entities that represent data assets and operation that no longer exist on the cluster. The purge functionality is available to users with administrator privilege; run a REST API command that lists one or more GUID values for the deleted entities.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

Enhancements to Basic Search in Atlas

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

System attributes filter searches

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



For information on using system attributes in Advanced Search, see Apache Atlas metadata attributes.

What's New in Cloud Connectors

Learn about the new features of Cloud Connectors in Cloudera Runtime 7.3.1.

Migration to AWS V2 SDK

The following improvements and enhancements have been introduced for AWS V2 SDK migration:

- Dual-layer server-side encryption (DSSE) has been enabled with AWS KMS keys
- AWS SDK V2 has been upgraded to 2.25.53

What's New in Apache Hive

There are no new features for Apache Hive in Cloudera Runtime 7.3.1.

What's New in Hue

Learn about the new features of Hue in Cloudera Runtime 7.3.1.

General availability (GA) of the SQL AI Assistant

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, promoting efficient and accurate practices for accessing and manipulating data. 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.

- To learn more about the supported models and services, limitations, and what data is shared with the LLMs, see About the Hue SQL AI Assistant.
- To set up and enable the SQL AI Assistant, see About setting up the Hue SQL AI Assistant.
- To see how to generate, edit, explain, optimize, and fix queries, see Starting the SQL AI Assistant in Hue.

Hue supports Python 3.9 on RHEL 8.8 and RHEL 8.10

Starting from the 7.3.1 release, Hue supports only Python 3.9 for RHEL 8.8 and RHEL 8.10. Before upgrading to CDP runtime 7.3.1, you must install Python 3.9 on all the Hue servers, as Hue requires a Python 3.9 version and does not start without it. For information about migrating from Python 3.8 to Python 3.9, see Migrating from Python 3.8 to Python 3.9 on RHEL 8.8 or RHEl 8.10.

What's New in Apache Iceberg

There are no new features for Apache Iceberg in Cloudera Runtime 7.3.1. You can continue to use the Iceberg capabilities that were available in the earlier version.

What's New in Apache Impala

Learn about the new features of Apache Impala in Cloudera Runtime 7.3.1.

Enhanced event processing information in /events web UI

The /events page of Catalogd now includes enhanced metrics, such as event processing lag and details on the current event batch. Error messages are highlighted at the top, if event processing stops. They will disappear after global INVALIDATE METADATA.

Apache Jira: IMPALA-12782

Improved query timeline with disk, network, and memory usage metrics

This update enhances the query timeline in the WebUI by adding disk and network usage metrics alongside CPU utilization. Metrics now display in human-readable formats like KB, MB, and GB. The update also introduces resizable and closable charts, zoom controls for easy navigation, and auto-scaling for timeticks during horizontal zoom. This enhanced display makes monitoring resource usage more intuitive for users.

Apache Jira: IMPALA-12364

Unicode column name support in Impala

Impala now supports Unicode characters in column names, aligning with Hive's support for non-ASCII characters. This enhancement leverages Hive's validateColumnName() function, which removes restrictions on column names at the metadata level. With this update, Impala allows greater flexibility for column naming while remaining consistent with Hive's metadata validation standards.

Apache Jira: IMPALA-12465

Support custom hash partitions at range level in Kudu tables

Impala now supports specifying custom hash partitions at the range level in Kudu tables. You can define hash schemas within specific partitions using the updated CREATE TABLE and ALTER TABLE syntax, and view them with the new SHOW HASH SCHEMA statement. This update aligns hash partitioning more closely with range partitioning, enhancing flexibility while maintaining backward compatibility.

Apache Jira: IMPALA-11430

What's New in Apache Kafka

Learn about the new features of Apache Kafka in Cloudera Runtime 7.3.1.

Kafka Rolling Restart check—all partitions fully replicated

A new broker rolling restart check option, all partitions fully replicated has been introduced. Selecting this option ensures that all partitions are in a fully synchronized state when a broker is stopped. For more information, see Rolling restart checks.

What's New in Livy

There are no new features for Livy in Cloudera Runtime 7.3.1.

What's New in Apache Oozie

Learn about the new features of Apache Oozie in Cloudera Runtime 7.3.1.

OpenJPA 3 upgrade in Apache Oozie

Third-party library OpenJPA is upgraded from version 2 to version 3 in Apache Oozie. This upgrade includes the following updates:

- New configuration properties
- Deprecated configuration properties
- Enhanced error handling

For more information, see OpenJPA upgrade.

What's New in Apache Ranger

Learn about the new features of Apache Ranger in Cloudera Runtime 7.3.1.

ZooKeeper SSL/TLS support for Ranger

Ranger and Ranger plugin audit to Solr supports ZooKeeper-SSL enabled connection.

Support multiple columns policy creation in Ranger for Grant/Revoke request

This enhancement supports multiple columns policy creation in Ranger for Grant/Revoke requests for Impala.

Ranger REST API improvements

Ranger REST APIs have the following changes:

- The following APIs have been removed:
  - assets/credstores - GET, POST, PUT
  - credstores/count - GET
  - credstores/{id} - GET
  - /xusers/auditmaps - GET
  - /xusers/auditmaps/count - GET
  - /xusers/permmaps - GET
  - /resource/{id} - GET
  - assets/policyList/{repository}
  - /groupgroups/* (All methods)
- The following APIs were not returning any access code when request is denied; now they suppose to 403:
  - /tags/tags
  - /tags/types
  - /tags/resources APIs
- Earlier When a non admin user makes a DELETE request to below endpoint, it was returning 405 method not allowed. However, now it returns 403.
  - /assets/resources/{resource_id}
- Earlier the API was not accessible for the keyadmin role users, but now it shall be accessible.
  - /xaudit/trx_log
- Earlier the below mentioned API was returning {OWNER} and {USER} users in the response but now onwards it will not return because access to the users list will be based on which role user is having permissions to which role user.
  - /service/xusers/users
- The API endpoint /xaudit/trx_log/{trx_log_id} was not accessible by keyadmin users. keyadmin users can access the transaction logs using the endpoint /xaudit/trx_log, hence, the keyadmin users should also be allowed to access the endpoint /xaudit/trx_log/{trx_log_id} for transaction log ids related to KMS audits.

What's New in Schema Registry

Learn about the new features of Schema Registry in Cloudera Runtime 7.3.1.

Enable SMM principal as trusted proxy user in Schema Registry

SMM usually connects to Schema Registry on behalf of an end user. For requests coming from SMM, Schema Registry can now extract and authorize the end user to authorize the request.

What's New in Apache Solr

Learn about the new features of Apache Solr in Cloudera Runtime 7.3.1.

Data Discovery and Exploration (Technical Preview)

A new Data Discovery and Exploration cluster definition is available in Data Hub. It lets you explore and discover data sets ad-hoc; doing relevance-based analytics over unstructured data (logs, images, text, PDFs, etc). The cluster definition deploys HDFS, Hue, Solr, Spark, Yarn, and ZooKeeper services. The cluster definition is available for AWS.

What's New in Apache Spark

Learn about the new features of Apache Spark in Cloudera Runtime 7.3.1.

Spark 2 removed from Cloudera Runtime

Spark 3 is the default Spark version in Cloudera Runtime. Spark 2 has been removed and no longer available in 7.3.1.

⚠ **Important:**

Spark 3 contains a large number of changes from Spark 2.

Refer to *Upgrading Spark* for more information on upgrading Spark clusters to 7.3.1.0, and *Migrating Spark Applications* for more information on migrating your existing Spark applications between versions 2 and 3.

Related Information

Upgrading Spark

Migrating Spark Applications

What's new in Streams Messaging Manager

Learn about the new features of Streams Messaging Manager (SMM) in Cloudera Runtime 7.3.1.

Validation for duplicate property keys in Kafka Connect connector configuration

When validating Kafka Connect connector configurations, a warning is displayed if the configuration contains duplicate property keys. Duplicate property keys are highlighted with orange. The form can still be validated with the warnings present, but if there are duplicates, you are notified that only the value of the last occurrence is used.

Search supports regular expressions

The search component on the Topics, Brokers, Consumers, Producers page can now perform a regexp search.

Visual clue when restarting on Kafka Connect

When clicking restart on Kafka Connect tasks or connectors, a loading circle is displayed in case of synchronous calls. The loading circle disappears once a response is received. For asynchronous calls, a pop-up is displayed, stating that the task or connector is restarted.

UX improvements

- Fixed text overflow in the side panel column headers
- Listing page table headers are now sticky of the nested table headers
- Listing page table styling has been improved for readability
- Filter selector drop-downs are now styled consistently
- Sidebar menu pop-ups are no longer hidden under tables
- Class names on the Kafka Connect popup are now wrapped into the containing pop-ups
- The password field is no longer obfuscated when using a file provider as a password
- Fixed the alignment of values on the Connector metrics page
- Source and sink connectors are now separate tabs on the connector creation modal
- Fixed visual issues on the topic creation modal
- Increased consistency in element contrast and text style throughout the UI
- Active and Inactive statuses now have high contrast
- The expand icon is now consistent throughout the UI

Expand security-related headers set by SMM

The following security related headers were added to SMM UI endpoints:

- Referrer-Policy
- Cross-Origin-Embedder-Policy
- Cross-Origin-Opener-Policy
- Cross-Origin-Resource-Policy

SMM uses trusted proxy authentication when connecting to Schema Registry

You can only interact with schemas through SMM if the necessary Ranger policies are set up for Schema Registry. For SMM UI, you must have the correct permissions to check messages deserialized with Avro on Data Explorer.

What's New in Apache Hadoop YARN

Learn about the new features of Apache Hadoop Yarn in Cloudera Runtime 7.3.1.

Queue Manager

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

For more information about Queue Manager, see Manage Queues.

FPGA as a resource type

You can use FPGA as a resource type. For more information, see Use FPGA scheduling.

New configuration property to enable or disable the YARN recommendation engine APIs

The YARN Recommendation API now recommends scaling cluster nodes up or down based on the demand and idle state of cluster resources. This feature can be turned on/off using the YARN configuration property yarn.cluster.scaling.recommendation.enable.

Unaffected Components in this release

There are no new features for the following components in Cloudera Runtime 7.3.1.

- Avro
- Cruise Control
- Hadoop
- HBase
- Hive
- HDFS
- Iceberg
- Knox
- Kudu
- Livy
- MapReduce
- Navigator Encrypt
- Ozone

- Phoenix
- Parquet
- Ranger KMS
- Sqoop
- Streams Replication Manager (SRM)
- Tez
- Zookeeper

What's new in Platform Support

You must be aware of the platform support for the Cloudera Runtime 7.3.1 release.

This section describes the platform support changes for the Cloudera Runtime 7.3.1 associated with Cloudera Public Cloud 7.3.1.

Platform Support Enhancements

- Default for all new environments:
  - OS support: RHEL 8.10
  - JDK version: JDK 17
  - Database version: PostgreSQL 14
  - Python version: 3.9
- No longer supported:
  - JDK 11 no longer supported (removed)

Cloudera Runtime Component Versions

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

Apache Components

Component	Version
Apache Arrow	0.11.1.7.3.1.0-197
Apache Atlas	2.1.0.7.3.1.0-197
Apache Calcite	1.25.0.7.3.1.0-197
Apache Avatica	1.22.0.7.3.1.0-197
Apache Avro	1.11.1.7.3.1.0-197
Apache Hadoop (Includes YARN and HDFS)	3.1.1.7.3.1.0-197
Apache HBase	2.4.17.7.3.1.0-197
Apache Flink	1.19.1.14.0.0
Apache Hive	3.1.3000.7.3.1.0-197
Apache Iceberg	1.3.1.7.3.1.0-197
Apache Impala	4.0.0.7.3.1.0-197
Apache Kafka	3.4.1.7.3.1.0-197
Apache Knox	2.0.0.7.3.1.0-197

Component	Version
Apache Kudu	1.17.0.7.3.1.0-197
Apache Livy	0.7.23000.7.3.1.0-197
Apache MapReduce	3.1.1.7.3.1.0-197
Apache NiFi	1.28.1.2.2.9.0
Apache NiFi Registry	1.28.1.2.2.9.0
Apache NiFi [Technical Preview]	2.0.0.4.2.1.0
Apache NiFi Registry [Technical Preview]	2.0.0.4.2.1.0
Apache Oozie	5.1.0.7.3.1.0-197
Apache ORC	1.8.3.7.3.1.0-197
Apache Parquet	1.12.3.7.3.1.0-197
Apache Phoenix	5.1.3.7.3.1.0-197
Apache Ranger	2.4.0.7.3.1.0-197
Apache Solr	8.11.2.7.3.1.0-197
Apache Spark	3.4.1.7.3.1.0-197
Apache Sqoop	1.4.7.7.3.1.0-197
Apache Tez	0.9.1.7.3.1.0-197
Apache ZooKeeper	3.8.1.7.3.1.0-197

Other Components

Component	Version
Cruise Control	2.5.116.7.3.1.0-197
Data Analytics Studio	1.4.2.7.3.1.0-197
GCS Connector	2.1.2.7.3.1.0-197
Hue	4.5.0.7.3.1.0-197
Search	1.0.0.7.3.1.0-197
Schema Registry	0.10.0.7.3.1.0-197
Streams Messaging Manager	2.3.0.7.3.1.0-197
Streams Replication Manager	1.1.0.7.3.1.0-197
Data Discovery and Exploration	Technical Preview

Connectors and Encryption Components

Component	Version
HBase connectors	1.0.0.7.3.1.0-197
Hive Meta Store (HMS)	1.0.0.7.3.1.0-197
Hive on Tez	1.0.0.7.3.1.0-197
Hive Warehouse Connector	1.0.0.7.3.1.0-197
Spark Atlas Connector	3.4.1.7.3.1.0-197
Spark Schema Registry	3.4.1.7.3.1.0-197

Note: Cloudera Ozone version 1.3.0 code is equivalent to Apache Ozone 1.4.0 in the 7.3.1 release. However, the version number will be reset in the next release.

Using the Cloudera Runtime Maven repository 7.3.1

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 7.3.1.0-197

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	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-aws-s3-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-azure-adls-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-classification-updater	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-client-common	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-client-v1	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-client-v2	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-client-v2-shaded	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-common	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-distro	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-docs	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-graphdb-api	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-graphdb-common	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-graphdb-janus	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-hdfs-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-index-repair-tool	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-intg	2.1.0.7.3.1.0-197

Project	groupId	artifactId	version
Atlas	org.apache.atlas	atlas-janusgraph-hbase2	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-notification	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-plugin-classloader	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-repository	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-server-api	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	atlas-testtools	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	hbase-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	hbase-bridge-shim	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	hbase-testing-util	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	hdfs-model	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	hive-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	hive-bridge-shim	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	impala-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	impala-bridge-shim	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	impala-hook-api	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	kafka-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	kafka-bridge-shim	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	navigator-to-atlas	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	sample-app	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	sqoop-bridge	2.1.0.7.3.1.0-197
Atlas	org.apache.atlas	sqoop-bridge-shim	2.1.0.7.3.1.0-197
Avro	org.apache.avro	avro	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-android	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-codegen-test	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-compiler	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-grpc	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-ipc	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-ipc-jetty	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-ipc-netty	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-mapred	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-maven-plugin	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-perf	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-protobuf	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-service-archetype	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-test-custom-conversions	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-thrift	1.11.1.7.3.1.0-197
Avro	org.apache.avro	avro-tools	1.11.1.7.3.1.0-197
Avro	org.apache.avro	trevni-avro	1.11.1.7.3.1.0-197
Avro	org.apache.avro	trevni-core	1.11.1.7.3.1.0-197

Project	groupId	artifactId	version
Calcite	org.apache.calcite	calcite-babel	1.25.0.7.3.1.0-197
Calcite	org.apache.calcite	calcite-core	1.25.0.7.3.1.0-197
Calcite	org.apache.calcite	calcite-druid	1.25.0.7.3.1.0-197
Calcite	org.apache.calcite	calcite-linq4j	1.25.0.7.3.1.0-197
Calcite	org.apache.calcite	calcite-server	1.25.0.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-aliyun	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-annotations	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-archive-logs	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-archives	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-assemblies	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-auth	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-aws	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-azure	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-azure-datalake	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-benchmark	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-build-tools	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-client	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-client-api	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-client-integration-tests	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-client-minicluster	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-client-runtime	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-cloud-storage	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-common	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-datajoin	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-distcp	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-extras	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-fs2img	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-gridmix	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-hdfs	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-hdfs-client	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-hdfs-httpfs	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-hdfs-native-client	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-hdfs-nfs	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-hdfs-rbf	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-kafka	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-kms	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-app	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-common	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-core	3.1.1.7.3.1.0-197

Project	groupId	artifactId	version
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-hs	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-hs-plugins	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-jobclient	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-nativetask	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-shuffle	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-client-uploader	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-mapreduce-examples	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-maven-plugins	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-minicluster	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-minikdc	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-nfs	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-openstack	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-resourceestimator	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-rumen	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-sls	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-streaming	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-tools-dist	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-api	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-applications-distributedshell	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-applications-unmanaged-am-launcher	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-client	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-common	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-registry	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-applicationhistoryservice	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-common	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-nodemanager	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-resourcemanager	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-router	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-sharedcachemanager	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-tests	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-timeline-pluginstorage	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-client	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-common	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-server-2	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-timelineservice-hbase-tests	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-server-web-proxy	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-services-api	3.1.1.7.3.1.0-197
Hadoop	org.apache.hadoop	hadoop-yarn-services-core	3.1.1.7.3.1.0-197

Project	groupId	artifactId	version
HBase	org.apache.hbase	hbase-annotations	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-asyncfs	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-checkstyle	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-client	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-client-project	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-common	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-endpoint	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-examples	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-external-blockcache	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-hadoop-compat	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-hadoop2-compat	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-hbtop	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-http	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-it	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-logging	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-mapreduce	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-metrics	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-metrics-api	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-procedure	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-protocol	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-protocol-shaded	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-replication	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-resource-bundle	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-rest	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-rsgroup	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-server	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-shaded-client	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-shaded-client-byo-hadoop	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-shaded-client-project	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-shaded-mapreduce	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-shaded-testing-util	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-shaded-testing-util-tester	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-shell	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-testing-util	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-thrift	2.4.17.7.3.1.0-197
HBase	org.apache.hbase	hbase-zookeeper	2.4.17.7.3.1.0-197
Hive	org.apache.hive	catalogd-unit	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-beeline	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-blobstore	3.1.3000.7.3.1.0-197

Project	groupId	artifactId	version
Hive	org.apache.hive	hive-classification	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-cli	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-common	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-contrib	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-exec	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-hbase-handler	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-hcatalog-it-unit	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-hplsql	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-iceberg-catalog	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-iceberg-handler	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-iceberg-shading	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-impala	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-custom-serde	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-iceberg	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-impala	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-minikdc	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-qfile	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-qfile-kudu	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-test-serde	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-unit	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-unit-hadoop2	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-it-util	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-jdbc	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-jdbc-handler	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-jmh	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-kudu-handler	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-llap-client	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-llap-common	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-llap-ext-client	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-llap-server	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-llap-tez	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-metastore	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-parser	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-pre-upgrade	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-serde	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-service	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-service-rpc	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-shims	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-standalone-metastore	3.1.3000.7.3.1.0-197

Project	groupId	artifactId	version
Hive	org.apache.hive	hive-storage-api	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-streaming	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-testutils	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-udf	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	hive-vector-code-gen	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	kafka-handler	3.1.3000.7.3.1.0-197
Hive	org.apache.hive	patched-iceberg-api	patched-1.3.1.7.3.1.0-197-3.1.3000.7.
Hive	org.apache.hive	patched-iceberg-core	patched-1.3.1.7.3.1.0-197-3.1.3000.7.
Hive Warehouse Connector	com.hortonworks.hive	hive-warehouse-connector-spark3_2.12	1.0.0.7.3.1.0-197
Kafka	org.apache.kafka	ci	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-api	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-basic-auth-extension	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-cloudera-authorization-extension	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-cloudera-common	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-cloudera-secret-storage	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-cloudera-security-policies	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-file	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-json	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-mirror	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-mirror-client	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-runtime	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	connect-transforms	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	generator	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-clients	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-cloudera-metrics-reporter_2.12	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-cloudera-metrics-reporter_2.13	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-cloudera-plugins	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-examples	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-group-coordinator	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-log4j-appender	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-metadata	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-raft	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-server-common	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-shell	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-storage	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-storage-api	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-examples	3.4.1.7.3.1.0-197

Project	groupId	artifactId	version
Kafka	org.apache.kafka	kafka-streams-scala_2.12	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-scala_2.13	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-test-utils	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0100	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0101	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0102	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-0110	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-10	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-11	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-20	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-21	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-22	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-23	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-24	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-25	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-26	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-27	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-28	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-30	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-31	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-32	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-streams-upgrade-system-tests-33	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka-tools	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka_2.12	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	kafka_2.13	3.4.1.7.3.1.0-197
Kafka	org.apache.kafka	trogdor	3.4.1.7.3.1.0-197
Knox	org.apache.knox	gateway-adapter	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-admin-ui	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-applications	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-cloud-bindings	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-demo-ldap	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-demo-ldap-launcher	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-discovery-ambari	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-discovery-cm	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-docker	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-i18n	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-i18n-logging-log4j	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-i18n-logging-sl4j	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-openapi-ui	2.0.0.7.3.1.0-197

Project	groupId	artifactId	version
Knox	org.apache.knox	gateway-performance-test	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-ha	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-identity-assertion-common	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-identity-assertion-concat	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-identity-assertion-hadoop-groups	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-identity-assertion-no-doas	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-identity-assertion-pseudo	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-identity-assertion-regex	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-identity-assertion-switchcase	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-jersey	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-rewrite	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-rewrite-common	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-rewrite-func-hostmap-static	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-rewrite-func-inbound-query-param	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-rewrite-func-service-registry	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-rewrite-step-encrypt-uri	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-rewrite-step-secure-query	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-authc-anon	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-authz-acls	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-authz-composite	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-authz-path-acls	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-clientcert	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-hadoopauth	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-jwt	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-pac4j	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-preauth	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-shiro	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-provider-security-webappsec	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-release	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-server	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-server-launcher	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-server-xforwarded-filter	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-admin	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-as	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-auth	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-definitions	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-hashicorp-vault	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-hbase	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-health	2.0.0.7.3.1.0-197

Project	groupId	artifactId	version
Knox	org.apache.knox	gateway-service-hive	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-idbroker	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-idbroker-plugins	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-impala	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-jkg	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-knoxsso	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-knoxssout	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-knoxtoken	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-livy	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-metadata	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-nifi	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-nifi-registry	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-remoteconfig	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-rm	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-session	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-storm	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-test	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-tgs	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-vault	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-service-webhdfs	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-shell	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-shell-launcher	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-shell-release	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-shell-samples	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-spi	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-spi-common	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-test	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-test-idbroker	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-test-release-utils	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-test-utils	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-topology-hadoop-xml	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-topology-simple	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-util-common	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-util-configinjector	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-util-launcher	2.0.0.7.3.1.0-197
Knox	org.apache.knox	gateway-util-urltemplate	2.0.0.7.3.1.0-197
Knox	org.apache.knox	hadoop-examples	2.0.0.7.3.1.0-197
Knox	org.apache.knox	knox-cli-launcher	2.0.0.7.3.1.0-197
Knox	org.apache.knox	knox-homepage-ui	2.0.0.7.3.1.0-197

Project	groupId	artifactId	version
Knox	org.apache.knox	knox-token-generation-ui	2.0.0.7.3.1.0-197
Knox	org.apache.knox	knox-token-management-ui	2.0.0.7.3.1.0-197
Knox	org.apache.knox	knox-webshell-ui	2.0.0.7.3.1.0-197
Knox	org.apache.knox	webhdfs-kerb-test	2.0.0.7.3.1.0-197
Knox	org.apache.knox	webhdfs-test	2.0.0.7.3.1.0-197
Kudu	org.apache.kudu	kudu-backup-tools	1.17.0.7.3.1.0-197
Kudu	org.apache.kudu	kudu-backup3_2.12	1.17.0.7.3.1.0-197
Kudu	org.apache.kudu	kudu-client	1.17.0.7.3.1.0-197
Kudu	org.apache.kudu	kudu-hive	1.17.0.7.3.1.0-197
Kudu	org.apache.kudu	kudu-spark3-tools_2.12	1.17.0.7.3.1.0-197
Kudu	org.apache.kudu	kudu-spark3_2.12	1.17.0.7.3.1.0-197
Kudu	org.apache.kudu	kudu-test-utils	1.17.0.7.3.1.0-197
Livy	org.apache.livy	livy-api	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-client-common	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-client-http	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-core_2.12	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-examples	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-integration-test	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-repl_2.12	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-rsc	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-scala-api_2.12	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-server	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-test-lib	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-thriftserver	0.7.23000.7.3.1.0-197
Livy	org.apache.livy	livy-thriftserver-session	0.7.23000.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-common	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-icu	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-kuromoji	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-morfologik	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-nori	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-opennlp	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-phonetic	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-smartcn	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-analyzers-stempel	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-backward-codecs	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-benchmark	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-classification	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-codecs	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-core	8.11.2.7.3.1.0-197

Project	groupId	artifactId	version
Lucene	org.apache.lucene	lucene-demo	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-expressions	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-facet	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-grouping	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-highlighter	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-join	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-memory	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-misc	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-monitor	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-queries	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-queryparser	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-replicator	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-sandbox	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-spatial-extras	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-spatial3d	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-suggest	8.11.2.7.3.1.0-197
Lucene	org.apache.lucene	lucene-test-framework	8.11.2.7.3.1.0-197
Oozie	org.apache.oozie	oozie-client	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-core	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-distro	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-examples	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-fluent-job-api	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-fluent-job-client	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-server	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-distcp	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-git	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-hcatalog	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-hive	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-hive2	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-oozie	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-spark3	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-sqoop	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-sharelib-streaming	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-tools	5.1.0.7.3.1.0-197
Oozie	org.apache.oozie	oozie-zookeeper-security-tests	5.1.0.7.3.1.0-197
ORC	org.apache.orc	orc-core	1.8.3.7.3.1.0-197
ORC	org.apache.orc	orc-examples	1.8.3.7.3.1.0-197
ORC	org.apache.orc	orc-mapreduce	1.8.3.7.3.1.0-197
ORC	org.apache.orc	orc-shims	1.8.3.7.3.1.0-197

Project	groupId	artifactId	version
ORC	org.apache.orc	orc-tools	1.8.3.7.3.1.0-197
Ozone	org.apache.ozone	hdds-annotation-processing	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-client	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-common	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-config	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-container-service	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-crypto-api	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-crypto-default	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-docs	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-erasurecode	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-hadoop-dependency-client	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-hadoop-dependency-server	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-hadoop-dependency-test	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-interface-admin	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-interface-client	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-interface-server	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-managed-rocksdb	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-rocks-native	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-server-framework	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-server-scm	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-test-utils	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	hdds-tools	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	mini-chaos-tests	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-client	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-common	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-csi	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-datanode	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-dist	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-filesystem	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-filesystem-common	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-filesystem-hadoop2	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-filesystem-hadoop3	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-filesystem-shaded	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-httpfsgateway	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-insight	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-integration-test	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-interface-client	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-interface-storage	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-manager	1.3.0.7.3.1.0-197

Project	groupId	artifactId	version
Ozone	org.apache.ozone	ozone-network-tests	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-recon	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-reconcodegen	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-s3-secret-store	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-s3gateway	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	ozone-tools	1.3.0.7.3.1.0-197
Ozone	org.apache.ozone	rocksdb-checkpoint-differ	1.3.0.7.3.1.0-197
Parquet	org.apache.parquet	parquet-avro	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-cascading	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-cascading3	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-column	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-common	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-encoding	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-format-structures	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-generator	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-hadoop	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-hadoop-bundle	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-jackson	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-pig	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-pig-bundle	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-protobuf	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-scala_2.12	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-thrift	1.12.3.7.3.1.0-197
Parquet	org.apache.parquet	parquet-tools	1.12.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-client-embedded-hbase-2.4	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-client-hbase-2.4	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-core	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.1.6	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.2.5	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.3.0	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.4.0	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.4.1	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.5.0	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-hbase-compat-2.5.4	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-pherf	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-queryserver	6.0.0.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-queryserver-client	6.0.0.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-queryserver-it	6.0.0.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-queryserver-load-balancer	6.0.0.7.3.1.0-197

Project	groupId	artifactId	version
Phoenix	org.apache.phoenix	phoenix-queryserver-orchestrator	6.0.0.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-server-hbase-2.4	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix-tracing-webapp	5.1.3.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix5-hive	6.0.0.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix5-hive-shaded	6.0.0.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix5-spark3	6.0.0.7.3.1.0-197
Phoenix	org.apache.phoenix	phoenix5-spark3-shaded	6.0.0.7.3.1.0-197
Ranger	org.apache.ranger	conditions-enrichers	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	credentialbuilder	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	embeddedwebserver	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	jisql	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ldapconfigcheck	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-adls-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-atlas-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-atlas-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-authn	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-common-ha	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-distro	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-examples-distro	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-gs-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-hbase-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-hbase-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-hdfs-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-hdfs-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-hive-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-hive-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-intg	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kafka-connect-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kafka-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kafka-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kms	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kms-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kms-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-knox-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-knox-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kudu-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kylin-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-kylin-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-metrics	2.4.0.7.3.1.0-197

Project	groupId	artifactId	version
Ranger	org.apache.ranger	ranger-nifi-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-nifi-registry-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-ozone-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-ozone-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-plugin-classloader	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-plugins-audit	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-plugins-common	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-plugins-cred	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-plugins-installer	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-policymigration	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-adls	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-chained-plugins	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-hook-abfs	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-hook-s3	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-intg	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-processor	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-s3	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-raz-s3-lib	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-rms-common	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-rms-hive	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-rms-plugins-common	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-rms-tools	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-rms-webapp	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-s3-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-sampleapp-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-schema-registry-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-solr-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-solr-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-sqoop-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-sqoop-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-storm-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-storm-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-tagsync	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-tools	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-trino-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-util	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-yarn-plugin	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ranger-yarn-plugin-shim	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	sample-client	2.4.0.7.3.1.0-197

Project	groupId	artifactId	version
Ranger	org.apache.ranger	sampleapp	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	shaded-raz-hook-abfs	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	shaded-raz-hook-s3	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	ugsync-util	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	unixauthclient	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	unixauthservice	2.4.0.7.3.1.0-197
Ranger	org.apache.ranger	unixusersync	2.4.0.7.3.1.0-197
Solr	org.apache.solr	solr-analysis-extras	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-analytics	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-cell	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-core	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-dataimporthandler	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-dataimporthandler-extras	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-gcs-repository	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-jaegertracer-configurator	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-langid	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-ltr	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-prometheus-exporter	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-s3-repository	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-security-util	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-solrj	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-test-framework	8.11.2.7.3.1.0-197
Solr	org.apache.solr	solr-velocity	8.11.2.7.3.1.0-197
Spark	org.apache.spark	spark-avro_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-catalyst_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-connect-client-jvm_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-connect-common_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-connect_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-core_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-graphx_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-hadoop-cloud_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-hive_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-kubernetes_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-kvstore_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-launcher_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-mllib-local_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-mllib_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-network-common_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-network-shuffle_2.12	3.4.1.7.3.1.0-197

Project	groupId	artifactId	version
Spark	org.apache.spark	spark-network-yarn_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-protobuf_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-repl_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-shaded-raz	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-sketch_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-sql-kafka-0-10_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-sql_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-streaming-kafka-0-10-assembly_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-streaming-kafka-0-10_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-streaming_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-tags_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-token-provider-kafka-0-10_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-unsafe_2.12	3.4.1.7.3.1.0-197
Spark	org.apache.spark	spark-yarn_2.12	3.4.1.7.3.1.0-197
Sqoop	org.apache.sqoop	sqoop	1.4.7.7.3.1.0-197
Sqoop	org.apache.sqoop	sqoop-test	1.4.7.7.3.1.0-197
Tez	org.apache.tez	hadoop-shim	0.9.1.7.3.1.0-197
Tez	org.apache.tez	hadoop-shim-2.8	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-api	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-aux-services	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-common	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-dag	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-examples	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-ext-service-tests	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-history-parser	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-javadoc-tools	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-job-analyzer	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-mapreduce	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-protobuf-history-plugin	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-runtime-internals	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-runtime-library	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-tests	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-yarn-timeline-cache-plugin	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-yarn-timeline-history	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-yarn-timeline-history-with-acls	0.9.1.7.3.1.0-197
Tez	org.apache.tez	tez-yarn-timeline-history-with-fs	0.9.1.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-angular	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-display	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-interpreter	0.8.2.7.3.1.0-197

Project	groupId	artifactId	version
Zeppelin	org.apache.zeppelin	zeppelin-jdbc	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-jupyter	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-livy	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-markdown	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-server	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-shaded-raz	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-shell	0.8.2.7.3.1.0-197
Zeppelin	org.apache.zeppelin	zeppelin-zengine	0.8.2.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-fatjar	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-loggraph	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-rest	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-contrib-zooinspector	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-it	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-jute	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-prometheus-metrics	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-election	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-lock	3.8.1.7.3.1.0-197
ZooKeeper	org.apache.zookeeper	zookeeper-recipes-queue	3.8.1.7.3.1.0-197

Fixed Issues In Cloudera Runtime 7.3.1

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

Fixed Issues in Atlas

Review the list of Atlas issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-69962: fetchType as "incremental" does full export instead of "CONNECTED"

Earlier, the first incremental export operation performed on a target entity used to fetch all entities even if they were not related to the targeted entity. This affected the performance as it imports more data than what was expected.

Now, the first incremental export will only fetch the entities which are related to the target entity. Also, if the target entity is connected to a lineage, then only the immediately connected entities in the lineage will get exported and not the whole lineage.

CDPD-67654: [Atlas] [navigator2atlas] Status of deleted table is ACTIVE in Atlas after navigator2atlas migration

Deleted hive tables migrated via the Navigator to Atlas transition may shown as active in Apache Atlas. Changes done in the Nav2Atlas module to set the relationType as hive_table_storagedesc of relationship attribute table for evey entity of hive_storagedesc.

CDPD-72732: [UCL] Incorrect Atlas audits generated for updates with atlas.hook.hive.skip.dml.messages set to true/false in 7.3.0.1 CDP Private Cloud Base

The Apache Atlas property atlas.hook.hive.skip.dml.messages = true can be used to reduce the number of audits that are generated for any DML command executed over a hive entity.

The default value for hive.split.update is set to true in 7.3.1 causing two audits to be generated for one update command: one delete and one insert. This will impact Apache Atlas when atlas.hook.hive.skip.dml.messages = false (Atlas is processing Data Manipulation events) and atlas.entity.audit.differential = false (Atlas logs the full entity metadata during every update).

CDPD-71516: Temporarily disable the tasks tab on Entity Detail page

The Entity Detail page was showing "Something went wrong". This is occurring because on loading the Entity Detail page, an API call (`/api/atlas/admin/tasks`) is made to get all the tasks that are created when deferred actions features are enabled. The Entity Detail page task tab and task API will display in UI depending upon the server side property atlas.tasks.ui.tab.enabled. Initially, this is set to false. Therefore, temporarily the task tab on entity detail page in UI is disabled.

Apache Jira:ATLAS-4880

OPSAPS-64385: Atlas's client.auth.enabled configuration is not configurable

In customer environments where user certifications are required to authenticate to services, the Apache Atlas web UI will constantly prompt for certifications. To solve this, the client.auth.enabled parameter is set to true by default. If it is needed to set it false, then you need to override the setting from safety-valve with a configuration snippet. Once it set to false, then no more certificate prompts will be displayed.

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

The issue of existing ATLAS OPTS not working for JDK17 is fixed.

Fixed Issues in Apache Avro

There are no fixed issues for Apache Avro in Cloudera Runtime 7.3.1.

Apache patch information

None

Fixed issues in Cloud Connectors

Review the list of Cloud Connectors issues that are resolved in Cloudera Runtime 7.3.1.

Apache patch information

- HADOOP-18855 - Tuning and stabilization of Vector IO
- MAPREDUCE-7474 - Improve commit resilience and performance in Manifest Committer for ABFS

Fixed issues in Cruise Control

Review the list of Cruise Control issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-69978: Cruise Control capacity.py script fails on Python 3

The script querying the capacity information is now fully compatible with Python 3.

Fixed Issues in Hadoop

There are no fixed issues for Hadoop in Cloudera Runtime 7.3.1.

Fixed Issues in HDFS

Review the list of HDFS issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-71677: When you are upgrading from CDP Private Cloud Base 7.1.9 SP1 to CDP Private Cloud Base 7.3.1, upgrade-rollback execution fails during HDFS rollback due to missing directory.

This issue is now resolved. The HDFS meta upgrade command is executed by creating the previous directory due to which the rollback does not fail.

OPSAPS-71390: COD cluster creation is failing on INT and displays the Failed to create HDFS directory /tmp error.

This issue is now resolved. Export options for jdk17 is added now.

OPSAPS-71188: Modify default value of dfs_image_transfer_bandwidthPerSec from 0 to a feasible value to mitigate RPC latency in the namenode.

This issue is now resolved.

OPSAPS-58777: HDFS Directories are created with root as user.

This issue is now resolved by fixinf service.sdl.

CDPD-67823: Ranger RMS gives all permissions to the user through the Create permission.

This issue is now resolved. An additional check is added to ensure that the user attempting to alter any HDFS directory that maps to the Hive database is the owner of the Hive database.

Fixed Issues in HBase

Review the list of HBase issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-67520: JWT authentication expects [sub] claim in the payload

A JWT payload can have a custom claim for Subject/Principal instead of the standard sub claim.

You can set the hbase.security.oauth.jwt.token.principal.claim configuration property in Cloudera Manager under HBase Service Advanced Configuration Snippet (Safety Valve) for hbase-site.xml to define the custom Subject/Principal claim.

CDPD-66387: RegionServer should be aborted when WAL.sync throws TimeoutIOException

This fix adds additional logic for WAL.sync. If WAL.sync gets a timeout exception, HBase wraps TimeoutIOException as a special WALSyncTimeoutIOException. When the upper layer such as HRegion.doMiniBatchMutate called by HRegion.batchMutation catches this special exception, HBase aborts the region server.

Apache Jira: HBASE-27230

CDPD-65373: Make delay prefetch property dynamically configurable

This change allows you to dynamically configure the hbase.hfile.prefetch.delay property using the Cloudera Manager. You need to update the value and refresh the HBase service. The new value is applied to the HBase service automatically.

Apache Jira: HBASE-28292

CDPD-74494: JVM crashes intermittently on ARM64 machines

After noticing the JVM crashes in the HBase service that is based on arm64 architecture and uses JDK 17, the fix is applied that refactors the module and the large implementation function into multiple smaller functions. The issue was observed in a specific module that had a very large member function.

Apache Jira: HBASE-28206

CDPD-73118: Bucket cache validation fails after a rolling restart resulting in an empty bucket cache without running the prefetch operations

During the retrieval of bucket cache from persistence, it was observed that, if an exception, other than the IOException occurs, the exception is not logged, and also the retrieval thread exits leaving the bucket cache in an uninitialized state, leaving it unusable.

This change enables the retrieval thread to print all types of exceptions and also reinitializes the bucket cache and makes it reusable.

Fixed Issues in Hive

Review the list of Hive issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-57121: ThreadPoolExecutorWithOomHook handling OutOfMemoryError

The ThreadPoolExecutorWithOomHook wasn't effectively handling OutOfMemoryError when executing tasks, as the exception was wrapped in ExecutionException, making it harder to detect.

The issue was fixed by updating ThreadPoolExecutorWithOomHook to properly invoke OutOfMemoryError hooks and stop HiveServer2 when required.

Apache Jira: HIVE-24411, HIVE-26955, IMPALA-8518

Fixed Issues in Hue

Review the list of Hue issues that are resolved in Cloudera Runtime 7.3.1

CDPD-65034: Receiving Error " TypeError: 'NoneType' object is not callable" in TCLIService.py when custom headers are being set

When XSRF (Cross-Site Request Forgery) or CSRF (Cross-Site Request Forgery) is enabled in Hive or Impala, you might encounter the error "Error " TypeError: 'NoneType' object is not callable” in TCLIService.py. You can resolve this issue by upgrading to 7.1.9 SP1 CHF3 or 7.3.1 versions.

Fixed Issues in Impala

Review the list of Impala issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-69856: SIGSEGV crash while accessing query state from concurrent access during query execution

A crash can occur due to concurrent updates and reads of execution state, such as through the WebUI, during query processing.

Ensured atomic updates of execution state to prevent conflicts and crashes during concurrent operations.

Apache Jira: IMPALA-12747

Missing txnId in tableWriteIds Mapping during AllocWriteIdEvent Processing

Fix issue of txnId not being added to tableWriteIds mapping in Catalog.

Apache Jira:: IMPALA-12851

CDPD-73442: Resolution of potential deadlock

This fix addresses a deadlock issue in long-running sessions with an active idle_query_timeout, which caused new queries to hang and prevented existing queries from expiring.

Apache Jira: IMPALA-13313

CDPD-71288: Retry HMS fetch failures to keep event-processor active

Metastore event processor enters an error state due to failures in creating a MetaStoreClient or fetching events, which should be retriable instead.

This issue is now fixed.

Apache Jira: IMPALA-12561

Some local file descriptors not released when using remote spilling

The issue that occurred during remote spilling when writing spilled data to local buffers has been fixed. The disk space occupied by the file can now be reclaimed.

Apache Jira:: IMPALA-12681

Handle empty string in StringValue::LargestSmallerString

StringValue::LeastSmallerString() did not account for empty strings, causing potential exceptions by using an invalid length.

The function now checks if the string is empty and returns an empty string if so. The function was also renamed to LargestSmallerString() to clarify its purpose.

Apache Jira:: IMPALA-12478

CDPD-67493 : ALTER_PARTITION self-event detection for partitions created via INSERT

Fix for incorrect identification on self events.

Apache Jira:: IMPALA-12356

CDPD-67912 : Failures in processing AbortTxnEvent

Fixes event processing errors when write IDs of an AbortTxnEvent are cleaned up by the HMS cleaner housekeeping threads.

Apache Jira:: IMPALA-12827

CDPD-67493: Implicit invalidate metadata on event failures

Implicitly invalidates a table instead of resulting in an ERROR state during event processing.

Apache Jira:: IMPALA-12832

IMPALA-12831: HdfsTable.toMinimalTCatalogObject() failed by concurrent modification

Fix race condition when a table is being invalidated and updated concurrently.

Apache Jira:: IMPALA-12831

Release JNI array if DeserializeThriftMsg failed

Fix conditional JVM heap leak in array allocation on deserialization failures.

Apache Jira:: IMPALA-12969

NPE in executing RELOAD events

Fixes the possibility of encountering a NullPointerException when refreshing a partition that has just been dropped.

Apache Jira:: IMPALA-12969

Event processing without hms_event_incremental_refresh_transactional_table

Fix event processor, which is not synching file metadata for non-partitioned ACID tables when incremental refresh on transactional tables is turned off.

Note:

- This issue only occurs when hms_event_incremental_refresh_transactional_table is set to 'false'
- This issue occurs on non-partitioned tables. Partitioned tables are not affected.

Apache Jira:: IMPALA-12835

Fixed Issues in Apache Iceberg

Review the list of iceberg issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-48395: Upgrade the Parquet version to 1.12.3 for Hive

This fix upgrades the Parquet version for Hive to 1.12.3, which is the same Parquet version that is used for Iceberg.

Fixed Issues in Kafka

Review the list of Kafka issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-65649: ReplicaAlterLogDirs stuck with Offset mismatch for the future replica

This is a backported fix, see KAFKA-9087 for more information.

CDPD-66986: Mirrormaker 2 auto.offset.reset=latest not working

This is a backported fix, see KAFKA-13988 for more information.

OPSAPS-71258: Kafka, SRM, and SMM cannot process messages compressed with Zstd or Snappy is /tmp is mounted as noexec

The issue is fixed by using JVM flags that point to a different temporary folder for extracting the native library.

CDPD-71433: Connect logical type null values are not handled in AvroConnectTranslator

When the time.precision.mode property is set to connect for the Debezium connector, the connect logical types are used and null values are now handled.

OPSAPS-69481: Some Kafka Connect metrics missing from Cloudera Manager due to conflicting definitions

Cloudera Manager now registers the metrics kafka_connect_connector_task_metrics_batch_size_avg and kafka_connect_connector_task_metrics_batch_size_max correctly.

Fixed Issues in Kudu

Review the list of Kudu issues that are resolved in Cloudera Runtime 7.3.1

KUDU-3576: Fix the Connection timeout After Tablet Server Restart

In Kudu, if a Java client application maintains an open connection to a tablet server and the tablet server is restarted or encounters a network error, the client cannot re-establish communication with the tablet server even after it comes back online. The fix resolves the issue by updating the Kudu Java client.

Apache Jira: KUDU-3576

KUDU-3496: Support for SPNEGO dedicated keytab

Kudu now supports configuring a dedicated Spnego keytab.

Apache Jira: KUDU-3496

KUDU-3524: Fix crash when sending periodic keep-alive requests

The fix ensures that Kudu clients do not crash when sending keep-alive requests.

Apache Jira: KUDU-3524

KUDU-3497: Optimize leader rebalancer algorithm

Optimized the leader balancing algorithm to effectively handle corner cases detailed in the Jira.

Apache Jira: KUDU-3497

KUDU-3447: Limit the usage of network bandwidth of tablet copying

Two new flags are introduced in Kudu CLI tools to copy tablets kudu tablet copy_from_remote command to limit the speed of the copy task and avoid resource contention.

Apache Jira: KUDU-3447

KUDU-3353: Add an immutable attribute to column schema

Introduced Immutable column. It’s useful to represent a semantically constant entity.

Apache Jira: KUDU-3353

KUDU-3351: Add insert error count metrics in WriteResponsePB

Statistics on various write operations are now available via Kudu client API at the session level.

Apache Jira: KUDU-3351

KUDU-3526: Scanner should bound with a tserver in java client

The scanner in the Kudu Java client now binds with the Kudu tablet server. This prevents scanning failures that occur when scanning from the leader replica and leadership changes to a different replica.

Apache Jira: KUDU-3526

Fixed Issues in Apache Knox

Review the list of Knox issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-73275: HTTP 404 responses while Knox is redeploying topologies

While you were redeploying topologies, Knox returned HTTP 404 responses.

Knox no longer returns HTTP 404 responses during topology redeployment, but returns HTTP 503 instead.

CDPD-70313: KNOX did not send Authentication header on FIPS configuration

KNOX neither sent the authentication header nor hadoop.auth cookie that was why the SMM UI sent back the HTTP 401 response and set the "www-authenticate": "Negotiate" header. Because of this, the SMM UI was inaccessible through Knox.

This issue is fixed now.

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
```

CDPD-69305: /plugins/policies/importPoliciesFromFile API returns 500 service connectivity error through Knox Proxy

The fix imports large policy files using the Ranger importPoliciesFromFile API through Knox.

Apache patch information

- KNOX-3073
- KNOX-3058
- KNOX-3055
- KNOX-3054
- KNOX-3053
- KNOX-3052
- KNOX-3050
- KNOX-3049

- KNOX-3045
- KNOX-3040
- KNOX-3038
- KNOX-3037
- KNOX-3036
- KNOX-3029
- KNOX-3028
- KNOX-3026
- KNOX-3024
- KNOX-3023
- KNOX-3022
- KNOX-3020
- KNOX-3019
- KNOX-3018
- KNOX-3017
- KNOX-3016
- KNOX-3012
- KNOX-3007
- KNOX-3006
- KNOX-3005
- KNOX-3002
- KNOX-3001
- KNOX-3000
- KNOX-2994
- KNOX-2985
- KNOX-2983
- KNOX-2980
- KNOX-2979
- KNOX-2978
- KNOX-2976
- KNOX-2975
- KNOX-2974
- KNOX-2973
- KNOX-2972
- KNOX-2971
- KNOX-2970
- KNOX-2969
- KNOX-2968
- KNOX-2966
- KNOX-2963
- KNOX-2961
- KNOX-2960
- KNOX-2959
- KNOX-2958
- KNOX-2955
- KNOX-2951
- KNOX-2949
- KNOX-2948
- KNOX-2947
- KNOX-2946

- KNOX-2929
- KNOX-2896
- KNOX-2881

Fixed Issues in Apache Livy

Review the list of Livy issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-71873 - UCL | CKP4| livyfoo0 kms proxy user is not allowed to access HDFS in 7.3.1.0

In the kms-core.xml file, the Livy proxy user is taken from Livy for Spark 3's configuration in Cloudera Public Cloud version 7.3.1 and above.

CDPD-73324 - LIVY_FOR_SPARK3 goes into down with Invalid Keystore format error in FIPS cluster

Fixed an issue that caused LIVY_FOR_SPARK3 to go into a bad state with Invalid Keystore format error in a 7.3.1 FIPS cluster.

Apache patch information

None

Fixed Issues in Oozie

Review the list of Oozie issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-70422: Cannot enforce Oozie parameter oozie.http.hostname

A new property named oozie.http.hostname.override is now introduced to specify the interface that the Oozie Server must be using.

CDPD-71117: Oozie server does not pass action start time to action conf causes a restarting launcher doesn't find child apps

Whenever Yarn restarted the Oozie Launcher AM, Oozie could not find the previously started child jobs due to a missing original start timestamp from the Oozie Server. And the previously started child Jobs were not terminated when the Launcher AM was restarted. This issue is now resolved.

CDPD-48664: Retry mechanism anomaly in Oozie with High Availability enabled

There was an issue with the retry mechanism in Oozie when High Availability was enabled. This issue is now resolved.

CDPD-49745: Expand app_path column in *_JOBS tables to allow HDFS paths longer than 255 characters

The APP_PATH column now supports storing paths longer than 255 characters.

Fixed Issues in Apache Parquet

There are no fixed issues for Parquet in Cloudera Runtime 7.3.1.

Apache Patch Information

None

Fixed Issues in Phoenix

Review the list of Phoenix issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-71553: OMID TSO server role fails during cluster installation

High Availability (HA) configuration generation is based on the number of nodes instead of a parameter in OMID.

OPSAPS-70507: ZDU runtime upgrade from 719 CHF4 to 7.3.0.1 fails for OMID

The OMID update issue is fixed.

OPSAPS-69838: OMID cannot connect contact the ZooKeeper cluster

OMID now connects to a secure port of ZooKeeper if TLS is enabled. Earlier it failed to communicate because port unification was missing.

OPSAPS-68583: ZooKeeper SSL/TLS support for OMID

OMID supports a secure connection to ZooKeeper if AutoTLS and ZooKeeper TLS are enabled.

OPSAPS-57949: OMID integration to Cloudera Manager

OMID now supports integration into the Cloudera Manager.

Fixed Issues in Ranger

Review the list of Ranger issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-73663: RMS server threw ConcurrentModificationException

The original ConcurrentModificationException was likely thrown when the resource-mappings were modified in response to changes in the Hive metadata while they were being serialized for downloading to the NameNode (or secondary-namenode).

The fix is to create a shallow copy of resource-mappings before applying deltas which ensures that resource-mappings are not modified while they are being serialized for downloading to the NameNode.

CDPD-73326: Reduce memory needed to create Ranger policy engine

Ranger policy engine creates a RangerPolicyResourceMatcher object for every resource specified either in policy or in a tag association. PolicyResourceMatcher, for the services that have more than one level in their resource hierarchy, consists of RangerResourceMatcher objects for each level in the resource-level hierarchy for the resource. In many cases, this leads to creation of multiple RangerResourceMatchers with identical resource specification.

The fix for this issue avoids creation of multiple RangerResourceMatcher objects by maintaining a cache of them in the RangerPluginContext object associated with the Ranger policy engine, thereby reducing policy engine's memory needs.

CDPD-73144: Trie to support processing of evaluators during traversal

Ranger policy engine uses trie data structure to organize resources for faster retrieval of policies/tags/zones associated with a given resource. When a resource consists of multiple elements, like database/table/column, as many trie instances are consulted to retrieve policies/tags/zones associated with the resource. Such multi-trie retrieval can be optimized with a 2-pass traversal - first pass to get count and the second pass to get the actual objects. Trie data structure used in Ranger policy engine should be updated to support this optimization.

Now, Trie to support processing of evaluators during traversal is enhanced.

CDPD-73102: Access issues for s3 express buckets

Fixed S3 Express bucket access with RAZ enabled in all regions.

CDPD-72203: Users observing role change from ROLE_SYS_ADMIN to ROLE_USER

Fixes role reset (to USER role) for users in usersync paged requests to ranger-admin.

CDPD-71719: Ranger override policy was not working

Ranger override policy was not allowing the access even though all permissions were given to the user.

This fix ensures that once all of the requested accesses are successfully allowed by (possibly multiple) Ranger policies, the access evaluation terminates with access allowed as the result.

CDPD-70081: "Drop database cascade" resulted in dropping of a table on which the user did not have access

Drop database cascade failed if the user did not have access to one or more of the underlying tables. It deleted the tables the user had access to but not others which caused the database to be not dropped as well.

This issue is fixed now.

CDPD-69488: Upgrade failure due to NPE in PatchForUpdatingServiceDefJson_J10058

Patch upgrade error failure in non-default service-def is fixed now.

CDPD-69305: /plugins/policies/importPoliciesFromFile API returns 500 service connectivity error through Knox Proxy

The fix imports large policy files using the Ranger importPoliciesFromFile API through Knox.

CDPD-68921: Exclude flag not taking effect for Ozone key resource in Ranger policy

Fix for exclude flag not taking effect for Ozone key resource in Ranger policy has been added.

CDPD-68853: Create function and Drop function commands are not supported when Ranger plugin is enabled

Support for Create and Drop function commands in Ranger trino plugin has been added.

CDPD-68827: Alter materialized view command is not working when Ranger plugin is enabled

Added support for Alter materialized view command in Ranger trino plugin.

CDPD-68826: Refresh materialized view command is not working when Ranger plugin is enabled

Added support for Refresh materialized view command in Ranger trino plugin.

CDPD-68376: Enable policy and tag deltas for Ranger admin and plugins by default

Policy and tag deltas for Ranger admin and plugins are enabled by default.

CDPD-68238: Update operations are not supported when Ranger plugin is enabled

The fix enables support for the update statement in the Ranger Trino plugin.

CDPD-67823: Ranger RMS gives all permissions to the user through the Create permission

An additional check is now made to ensure that the user attempting to alter a HDFS directory that maps to the Hive database is owner of the Hive database for the attempted operation is allowed.

CDPD-67193: Issue with inactivityTimeout getting reset

The inactivityTimeout was getting reset when a user updated its profile from the UserProfile page.

Fixed issue of not resetting inactivityTimeout to a default value of 15 minutes when user updates its profile from UserProfile page on Ranger Admin UI.

CDPD-66842: Ranger Admin server gives empty response

Ranger Admin server gave an empty response when a user with user-role tried to update lastname or email address.

The issue is fixed now. Error response with message will be shown when a user with user-role tries to add/update last name or email address.

CDPD-66839: Enhance perf-tracer to get CPU time when possible

Ranger module is instrumented with performance measurement code. It enables performance logging for the module and helps in measuring the amount of time spent during execution of various methods/functions during its operation. For achieving more precise time measurement, this feature supports nanosecond precision when the JVM version supports it.

CDPD-66624: Transform URLs with or without “/” at the end issue

The fix enables the transformation step handle “/” at the end of the path.

CDPD-66404: Merging apache ranger jiras for handling local storage data for column show/hide functionality

Implemented Column Hide/Show functionality in Audit Plugin Status tab.

CDPD-66358: HS2 logs having a huge number of WARN logs

HS2 logs had a huge number of WARN logs from RangerHiveAuthorizer regarding connection to HMS for fetching Hive object owner.

This fix addresses the issue where HS2 logs have a huge number of WARN logs.

CDPD-66136: Display of query information for Show databases/schemas command on Ranger Admin UI

In Ranger React UI, if the resource type for certain commands were logged as "null" in the audits, then in the access audits, the information of the query/operations performed would not be displayed.

This ticket addresses the issue and displays the query/operation information for access audits where the resource type was "null".

CDPD-66092: Ranger Javapatch failure even if service-defs do not exist in Ranger DB

Added support to upgrade non-default service-defs in Ranger.

CDPD-65923: Audit logs for Mask and Row policy does not show policy condition under policy item

The fix now shows policy conditions under policy items for Mask and Row policy Audit logs.

CDPD-65650: Pagination missing on the Ranger Admin - Plugin Status page

This fix offers the following:

- Sorting works properly after this patch.
- Pagination added.

CDPD-63891: Backport the ranger-trino changes from upstream to downstream

Trino support in Ranger has been added.

OPSAPS-70838: Flink user should be add by default in ATLAS_HOOK topic policy in Ranger >> cm_kafka

The "flink" service user is granted publish access on the ATLAS_HOOK topic by default in the Kafka Ranger policy configuration.

OPSAPS-69411: Update AuthzMigrator GBN to point to latest non-expired GBN

Users will now be able to export sentry data only for given Hive objects (databases and tables and the respective URLs) by using the config "authorization.migration.export.migration_objects" during export.

OPSAPS-68252: "Ranger RMS Database Full Sync" option was not visible on mow-int cluster setup for hrt_qa user (7.13.0.0)

The fix makes the command visible on cloud clusters when the user has minimum EnvironmentAdmin privilege.

Apache Patch information

- RANGER-4973
- RANGER-4972
- RANGER-4960
- RANGER-4933
- RANGER-4912
- RANGER-4905
- RANGER-4893
- RANGER-4833
- RANGER-4823
- RANGER-4819

- RANGER-4818
- RANGER-4802
- RANGER-4799
- RANGER-4798
- RANGER-4797
- RANGER-4796
- RANGER-4791
- RANGER-4786
- RANGER-4782
- RANGER-4781
- RANGER-4780
- RANGER-4774
- RANGER-4767
- RANGER-4753
- RANGER-4747
- RANGER-4745
- RANGER-4737
- RANGER-4729
- RANGER-4722
- RANGER-4720
- RANGER-4718
- RANGER-4717
- RANGER-4710
- RANGER-4699
- RANGER-4698
- RANGER-4690
- RANGER-4689
- RANGER-4688
- RANGER-4681
- RANGER-4673
- RANGER-4668
- RANGER-4653
- RANGER-4641
- RANGER-4611
- RANGER-4609
- RANGER-4607
- RANGER-4598
- RANGER-4597
- RANGER-4596
- RANGER-4595
- RANGER-4594
- RANGER-4593
- RANGER-4591
- RANGER-4590
- RANGER-4589
- RANGER-4588
- RANGER-4586
- RANGER-4578
- RANGER-4577
- RANGER-4576

- RANGER-4575
- RANGER-4574
- RANGER-4573
- RANGER-4568
- RANGER-4555
- RANGER-4554
- RANGER-4553
- RANGER-4552
- RANGER-4551
- RANGER-4550
- RANGER-4549
- RANGER-4548
- RANGER-4547
- RANGER-4546
- RANGER-4545
- RANGER-4544
- RANGER-4532
- RANGER-4515
- RANGER-4513
- RANGER-4492
- RANGER-4370
- RANGER-4303
- RANGER-4278
- RANGER-4261
- RANGER-4229
- RANGER-4221
- RANGER-4172
- RANGER-4010
- RANGER-3805
- RANGER-3772
- RANGER-3759
- RANGER-3745
- RANGER-3657
- RANGER-3182
- RANGER-3174

Fixed Issues in Schema Registry

Review the list of Schema Registry issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-68708: Schema Registry might fail to start if a load balancer address is specified in Ranger

Schema Registry now always ensures that the address it uses to connect to Ranger ends with a trailing slash (/). As a result, Schema Registry no longer fails to start if Ranger has a load balancer address configured that does not end with a trailing slash.

Fixed Issues in Apache Solr

Review the list of Solr issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-71690: Update control group V2 configuration parameters

The default values of the control group (CGroup) V2 configuration parameters are updated in Cloudera Manager for the Solr service. The following table describes the default values of the corresponding V2 parameters.

Parameter name	Default values
memory.high	-1
memory.max	-1
io.weight	100
cpu.weight	100

For more information on CGroup V2 parameters, see Configuring Resource Parameters.

Fixed Issues in Spark

Review the list of Spark issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-74697 - Spark Iceberg vectorized Parquet read of decimal column is incorrect

CDPD-72774 - Use common versions of commons-dbcp2 and commons-pool2

CDPD-70114 - Redirect spark-submit, spark-shell etc. scripts to their Spark 3 counterparts

CDPD-58844 - Spark - Upgrade Janino to 3.1.10 due to CVE-2023-33546

CDPD-48171 - Spark3 - Upgrade snakeyaml due to CVE-2022-1471

Apache patch information

None

Fixed Issues in Streams Messaging Manager

Review the list of Streams Messaging Manager (SMM) issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-71258: Kafka, SRM, and SMM cannot process messages compressed with Zstd or Snappy if /tmp is mounted as noexec

The issue is fixed by using JVM flags that point to a different temporary folder for extracting the native library.

CDPD-72543: Security headers are not set for static files in SMM

SMM now applies the following security-related headers to static files:

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

CDPD-73643: Unused CM_USER parameter is visible in /cm-configs internal endpoint

The unused CM_USER field has been removed from the /cm-configs internal endpoint

CDPD-70313: KNOX does not send Authentication header on FIPS configuration

KNOX now sends the Authentication header on FIPS clusters.

Fixed Issues in Streams Replication Manager

Review the list of Streams Replication Manager (SRM) issues that are resolved in Cloudera Runtime 7.3.1.

OPSAPS-71258: Kafka, SRM, and SMM cannot process messages compressed with Zstd or Snappy if /tmp is mounted as noexec

The issue is fixed by using JVM flags that point to a different temporary folder for extracting the native library.

Fixed Issues in Apache Tez

There are no fixed issues for Tez in Cloudera Runtime 7.3.1.

Fixed Issues in YARN and YARN Queue Manager

Review the list of YARN and YARN Queue Manager issues that are resolved in Cloudera Runtime 7.3.1.

COMPX-17702: Backport - YARN-10345 - HsWebServices containerlogs does not honor ACLs for completed jobs

The following rest APIs now have ACL authorization:

- /ws/v1/history/containerlogs/{containerid}/{filename}
- /ws/v1/history/containers/{containerid}/logs

Apache Jira: YARN-10345

COMPX-16285: Optimize system credentials sent in node heartbeat responses

Previously, the heartbeat responses set all application's tokens even though all applications were not active on a node. Hence, for each node and each heartbeat too many SystemCredentialsForAppsProto objects were created. This issue is now resolved and the system credentials sent in node heartbeat responses are optimized..

Apache Jira: YARN-6523

CDPD-73754: Yarn Application Master Node web link is broken on yarnuiv2 page

Previously, the RM did not open the Yarn application manager node web link on the **yarnuiv2** page because the URL ended with a /. This issue is now resolved and the last character / is now removed from the URL.

Apache Jira: YARN-11729

Fixed Issues in Zookeeper

Review the list of ZooKeeper issues that are resolved in Cloudera Runtime 7.3.1.

CDPD-67821: Zookeeper - Information disclosure in persistent watcher handling(CVE-2024-23944)

There was information disclosure in persistent watchers handling in Apache ZooKeeper due to CVE-2024-23944. This issue is now fixed.

Apache Jira: ZOOKEEPER-4799

CDPD-66977: Backport ZOOKEEPER-4804 Use daemon threads for Netty client

Previously, when the Netty client was used, the Java process did not respond on System.exit when the Zookeeper connection was open. This issue was caused by the non-daemon threads created by Netty. This issue is now resolved.

Apache Jira: ZOOKEEPER-4804

Known Issues In Cloudera Runtime 7.3.1

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

Known Issues in Apache Atlas

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

CDPD-67450: Table name renaming operation is not updating or creating iceberg_table entity

Renaming an Iceberg table does not update the corresponding Atlas entity.

None.

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

None

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

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

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 ALTERTABLE_* lineage are missing for Spark tables created using spark3-shell

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

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

The ALTERTABLE_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-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

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

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-11692: Navigator table creation time not converted to Atlas

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

None

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

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

None

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

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

None

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

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

None

CDPD-9095: Duplicate audits for renaming Hive tables

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

None

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

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

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

None

CDPD-7781: TLS certificates not validated on Firefox

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

None

CDPD-6675: Irregular qualifiedName format for Azure storage

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

None

CDPD-4762: Spark metadata order may affect lineage

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

None

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

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

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

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

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

None

CDPD-3160: Hive lineage missing for INSERT OVERWRITE queries

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

None

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

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

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

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

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

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

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

The free text search bar in the top of the screen allows you to search across entity types and through all text attributes for all entities. The search shows the top 5 results that match the search terms at any place in the text (*term* logic). It also shows suggestions that match the search terms that begin with the term (term* logic). However, in this release, the search results are case-sensitive.

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

CDPD-1823: Queries with ? wildcard return unexpected results

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

None

CDPD-1664: Guest users are redirected incorrectly

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

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

CDPD-922: IsUnique relationship attribute not honored

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

None

CDPD-65619: Newly created Iceberg tables do not show up under hive_db entity

Currently, on single typename is shown under the Tables tab. Both Iceberg and Hive tables cannot be shown when they are created in the same hive_db entity.

CDPD-75994: Post DL regular upgrade (non ZDU) to 7.3.1, "Exception in getKafkaConsumer ,WakeupException: null" is seen

After the data lake is upgraded to 7.3.1, sometimes Atlas Hook does not function when Apache Atlas and Apache Kafka are started at the same time, thus Atlas is unable to connect to Kafka while Kafka is still being set up. Atlas performs only three attempts.

Restart the cluster, after the upgrade to trigger to reconnect to Apache Kafka. The Kafka consumer creation should be retried if the Kafka service is unavailable during Atlas startup.

CDPD-76035: Resource lookup for Atlas service is failing

Once the Atlas configuration snippet atlas.authentication.method.file is enabled and a classification is created, these do not synchronize correctly to the Type Category resource field setting of Apache Ranger. The newly created classification won't be able to be selected as the Type Name.

CDPD-74180: Export/Import : If Shell entities have a lineage, it is not exported

If there is a shell entity which has lineage, while using the Export API, that shell entity will not be exported in the zip file.

CDPD-66938: [Analyze] [Atlas] [FIPS] test_time_range tests fail

When the Apache Atlas server is running on a node which has time zone other than UTC, there might be a time of day when the search results might differ if the relative **CreateTime** date range filters of TODAY, YESTERDAY, etc. are used.

Use explicit date range filters instead of using relative date range filters, such as, TODAY, YESTERDAY.

CDPD-70321: Atlas Parallel import is failing with various errors

During a parallel import-export activity with six iceberg table policies with exportOption as db1.* for all six exports, all import fail after the exports.

CDPD-76269: POST Rolling Upgrade performed from 7.1.7.3000 to 7.3.1.0 , and then downgraded from 7.3.1.0 to 7.1.7.3000 , updating edge to enable tag propagation is failing

When performing a rolling upgrade from 7.1.7.3000 to 7.3.1.0, and then downgrading from 7.3.1.0 to 7.1.7.3000, updating edge to enable tag propagation fails.

Tag propagation works well after a restart.

Known Issues in Apache Avro

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

CDPD-23451: Avro library depends on the already EOL jackson-mapper-asl 1.9.13-cloudera.1 that also contains a couple of CVEs. The jackson library is part of the Avro API so cannot be changed without a complete rebase.

None.

Known Issues in Cloud Connectors

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

AWS SDK 2.25.53 warning about transfer manager not using CRT client

Due to the AWS SDK 2.25.53 upgrade, the following warning might be seen:

```
5645:2024-09-13 16:29:17,375 [setup] WARN  s3.S3TransferManager
          (LoggerAdapter.java:warn(225)) - The provided S3AsyncC
lient is an instance of
          MultipartS3AsyncClient, and thus multipart download fe
ature is not enabled. To benefit
          from all features, consider using S3AsyncClient.crtBu
ilder().build() instead
```

This error message is completely harmless and should be ignored. For more information, see HADOOP-19272.

None

Known issues in Cruise Control

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

CDPD-44676: Rebalancing with Cruise Control does not work if the metric reporter fails 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.CruiseCont
rolMetricsReporter:
       [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 moving the host to a different cluster. For more information, see Moving a Host Between Clusters

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

Known Issues in Apache Hadoop

There are no known issues for Hadoop in Cloudera Runtime 7.3.1.

Known Issues in Apache HBase

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

CDPD-60862: Rolling restart fails during ZDU when DDL operations are in progress

During a Zero Downtime Upgrade (ZDU), the rolling restart of services that support Data Definition Language (DDL) statements might fail if DDL operations are in progress during the upgrade. As a result, ensure that you do not run DDL statements during ZDU.

The following services support DDL statements:

- Impala
- Hive – using HiveQL
- Spark – using SparkSQL
- HBase
- Phoenix
- Kafka

Data Manipulation Lanaguage (DML) statements are not impacted and can be used during ZDU. Following the successful upgrade, you can resume running DDL statements.

None. Cloudera recommends modifying applications to not use DDL statements for the duration of the upgrade. If the upgrade is already in progress, and you have experienced a service failure, you can remove the DDLs in-flight and resume the upgrade from the point of failure.

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

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

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

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

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

HDFS encryption with HBase

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

Workaround: N/A

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

AccessController postOperation problems in asynchronous operations

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

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

Workaround: N/A

Apache Issue: HBASE-6992

HBase shutdown can lead to inconsistencies in META

Cloudera Manager uses an incorrect shutdown command. This prevents graceful shutdown of the HBase service and forces Cloudera Manager to kill the processes instead. It can lead to inconsistencies in Meta.

Workaround: Run the following command instead of shutting down the HBase service using Cloudera Manager.

```
hbase master stop --shutDownCluster
```

The command output must end with Closing master protocol: MasterService phrase. You can verify the command execution by checking the master logs. The log must contain Cluster shutdown requested of master=xxx and the closing of regions. Upon successful execution, the RegionServers start shutting down.

Note: The command does not stop the *REST Server* and the *Thrift Server* role instances. You can safely shut down them from Cloudera Manager later.

If you find any inconsistencies, please contact Cloudera Support.

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

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

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

Apache Issue: N/A

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

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

Workaround: N/A

Apache Issue: N/A

Known Issues in HDFS

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

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.

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-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 you have attempted to use all the datanodes in the cluster, and failed to write to each of them as they were restarted. This only happen on small clusters of less than 10 datanodes, because larger clusters have more spare nodes to allow the write to continue.

None.

CDPD-60873: java.io.IOException:Encountered "status=ERROR, status message, ack with firstBadLink" while fixing the HDFS corrupt file during rollback.

Increase the value of dfs.client.block.write.retries to the number of nodes in the cluster and perform Deploy client configuration procedure for rectification.

CDPD-60431: Configuration difference between 7.1.7 SP2 and 7.1.9.0 results

Component	Configuration	Old Value	New Value	Description
HDFS	dfs.permissions.ContentSummary.subAccess	Not Set	True	Performance optimization for NameNode content summary API
HDFS	dfs.datanode.handler.count	3	10	Optimal value for DN server threads on large clusters

None.

CDPD-60387: Configuration difference between 7.1.8.3 and 7.1.9.0 results

Component	Configuration	Old Value	New Value	Description
HDFS	dfs.namenode.accesstime.precision	Not set	0	Optimal value for NameNode performance on large clusters
HDFS	dfs.datanode.handler.count	3	10	Optimal value for DN server threads on large clusters

None.

OPSAPS-64307: When the JournalNodes on a cluster are restarted, the Add new NameNode wizard for HDFS service might fail to bootstrap the new NameNode. If there was no new fsImage created from the time JournalNodes restarted, during the restart the edit logs were rolled in the system.

If the bootstraping fails during the Add new NameNode wizard, then perform the following steps:

1. Delete the newly added NameNode and FailoverController
2. Move the active HDFS NameNode to safe mode
3. Perform the Save Namespace operation on the active HDFS NameNode
4. Leave safe mode on the active HDFS NameNode
5. Add the new NameNode again

Note: Entering safe mode disables writes to HDFS which causes a service disruption. If you cannot enter the safe mode, delete the newly added NameNode and FailoverController in the HDFS service and wait until HDFS automatically creates a new fsImage and then add the new NameNode again with the wizard.

OPSAPS-64363: Deleting of additional Standby Namenode does not delete the ZKFC role and this has to be done manually.

None.

CDPD-28390: Rolling restart of the HDFS JournalNodes may time out on Ubuntu20.

If the restart operation times out, you can manually stop and restart the Name Node and Journal Node services one by one.

OPSAPS-55788: WebHDFS is always enabled. The Enable WebHDFS option does not take effect.

None.

OPSAPS-63299: Disable HA command for a nameservice does not work if the nameservice has more than 2 NameNodes defined.

None.

OPSAPS-63301: Deleting nameservice command does not delete all the NameNodes belonging to the nameservice, if there are more than two NameNodes that are assigned to the nameservice.

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)
- NameNode Federation (HDFS-1052)

- NameNode Port-based Selective Encryption (HDFS-13541)
- Non-Volatile Storage Class Memory (SCM) in HDFS Cache Directives (HDFS-13762)
- OpenStack Swift (HADOOP-8545)
- SFTP FileSystem (HADOOP-5732)
- Storage policy satisfier (HDFS-10285)

Technical Service Bulletins

TSB 2022-549: Possible HDFS Erasure Coded (EC) data loss when EC blocks are over-replicated

Cloudera has detected a bug that can cause loss of data that is stored in HDFS Erasure Coded (EC) files in an unlikely scenario.

Some EC blocks may be inadvertently deleted due to a bug in how the NameNode chooses excess or over-replicated block replicas for deletion. One possible cause of over-replication is running the HDFS balancer soon after a NameNode goes into failover mode.

In a rare situation, the redundant blocks can be placed in such a way that one replica is in one rack, and few redundant replicas are in the same rack. Such placement causes a counting bug (HDFS-16420) to be triggered. Instead of deleting just the redundant replicas, the original replica may also be deleted.

Usually this is not an issue, because the lost replica can be detected and reconstructed from the remaining data and parity blocks. However, if multiple blocks in an EC Block Group are affected by this counting bug within a short time, the block cannot be reconstructed anymore. For example, 4 blocks are affected out of 9 for the RS(6,3) policy.

Another situation is recommissioning multiple nodes back into the same rack of the cluster where the current live replica exists.

Upstream JIRA

HDFS-16420

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: TSB 2022-549: Possible HDFS Erasure Coded (EC) data loss when EC blocks are over-replicated

Known Issues in Apache Hive

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

DAG not retried after failure

When executing a Hive query, if the ApplicationMaster container fails, Hive does not retry the DAG if the failure message contains some diagnostic information including a line break, leading to query failure (instead of retry).

None

Known Issues in Hue

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

CDPD-58978: Batch query execution using Hue fails with Kerberos error

When you run Impala queries in a batch mode, you enounter 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.

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.

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
```

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 in Hue configurations in Cloudera Manager:

```
[indexer]
enable_direct_upload=true
```

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.

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.

CDPD-56888: Renaming a folder with special characters results in a duplicate folder with a new name on AWS S3.

On AWS S3, if you try to rename a folder with special characters in its name, a new folder is created as a copy of the original folder with its contents. Also, you may not be able to delete the folder containing special characters.

You can rename or delete a directory having special characters using the HDFS commands as follows:

1. SSH into your CDP environment host.
2. To delete a directory within your S3 bucket, run the following command:

```
hdfs dfs -rm -r [***COMPLETE-PATH-TO-S3-BUCKET***]/[***DIREC
TORY-NAME***]
```

3. To rename a folder, create a new directory and run the following command to move files from the source directory to the target directory:

   ```
   hdfs dfs -mkdir [***DIRECTORY-NAME***]
   ```

   ```
   hdfs dfs -mv [***COMPLETE-PATH-TO-S3-BUCKET***]/[***SOURCE-D
   IRECTORY***] [***COMPLETE-PATH-TO-S3-BUCKET***]/[***TARGET-D
   IRECTORY***]
   ```

CDPD-48146: Error while browsing S3 buckets or ADLS containers from the left-assist panel

You may see the following error while trying to access the S3 buckets or ADLS containers from the left-assist panel in Hue: Failed to retrieve buckets: :1:0: syntax error.

Access the S3 buckets or ADLS containers using the File Browser.

CDPD-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
```

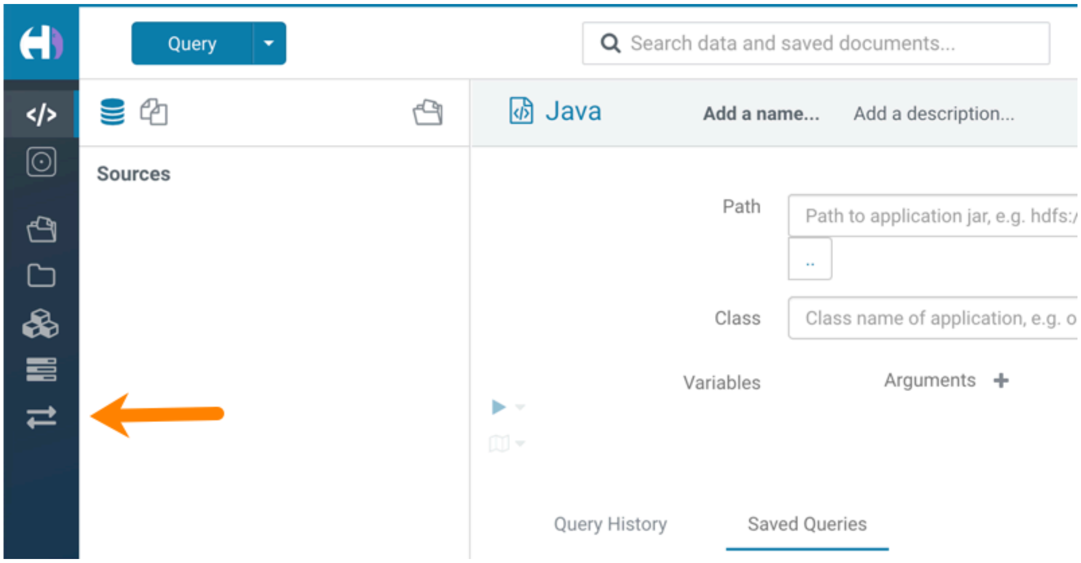
CDPD-42619: Unable to import a large CSV file from the local workstation

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.

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:



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.

Known Issues in Apache Iceberg

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

CDPD-75667: Querying an Iceberg table with a TIMESTAMP_LTZ column can result in data loss

When you query an Iceberg table that has a TIMESTAMP_LTZ column, the query could result in data loss.

When creating Iceberg tables from Spark, set the following Spark configuration to avoid creating columns with the TIMESTAMP_LTZ type:

```
spark.sql.timestampType=TIMESTAMP_NTZ
```

Apache JIRA: IMPALA-13484

CDPD-75088: Iceberg tables in azure cannot be partitioned by strings ending in '.'

In an Azure environment, you cannot create Iceberg tables from Spark that are partitioned by string columns having a partition value that contains the period (.) character. The query fails with the following error:

```
24/10/08 18:14:12 WARN  scheduler.TaskSetManager: [task-result-g
etter-2]: Lost task 0.0 in stage 2.0 (TID 2) (spark-sfvq0t-compu
te0.spark-r9.l2ov-m7vs.int.cldr.work executor 1): java.lang.Ille
galArgumentException: ABFS does not allow files or directories to
 end with a dot.
```

None.

CDPD-72942: Unable to read Iceberg table from Hive after writing data through Apache Flink

If you create an Iceberg table with default values using Hive and insert data into the table through Apache Flink, you cannot then read the Iceberg table from Hive using the Beeline client, and the query fails with the following error:

```
Error while compiling statement: java.io.IOException: java.io.IO
Exception: Cannot create an instance of InputFormat class org.ap
ache.hadoop.mapred.FileInputFormat as specified in mapredWork!
```

The issue persists even after you use the ALTER TABLE statement to set the engine.hive.enabled table property to "true".

None.

Apache JIRA: HIVE-28525

CDPD-71962: Hive cannot write to a Spark Iceberg table bucketed by date column

If you have used Spark to create an Iceberg table that is bucketed by the "date" column and then try inserting or updating this Iceberg table using Hive, the query fails with the following error:

```
Error: Error while compiling statement: FAILED: RuntimeException
 org.apache.hadoop.hive.ql.exec.UDFArgumentException:  ICEBERG_B
UCKET() only takes STRING/CHAR/VARCHAR/BINARY/INT/LONG/DECIMAL/F
LOAT/DOUBLE types as first argument, got DATE (state=42000,code=
40000)
```

This issue does not occur if the Iceberg table is created through Hive.

None.

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

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_REA D=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.

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_VER
SION AS OF <preTruncateSnapshotId>;
```

Known Issues in Apache Impala

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

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-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-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-1024: Impala BE cannot parse Avro schema that contains a trailing semi-colon

If an Avro table has a schema definition with a trailing semicolon, Impala encounters an error when the table is queried.

Remove trailing semicolon from the Avro schema.

IMPALA-1652: Incorrect results with basic predicate on CHAR typed column

When comparing a CHAR column value to a string literal, the literal value is not blank-padded and so the comparison might fail when it should match.

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

IMPALA-1821: Casting scenarios with invalid/inconsistent results

Using a CAST() function to convert large literal values to smaller types, or to convert special values such as NaN or Inf, produces values not consistent with other database systems. This could lead to unexpected results from queries.

None

IMPALA-2005: A failed CTAS does not drop the table if the insert fails

If a CREATE TABLE AS SELECT operation successfully creates the target table but an error occurs while querying the source table or copying the data, the new table is left behind rather than being dropped.

Drop the new table manually after a failed CREATE TABLE AS SELECT

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

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

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

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

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

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

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

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

None

IMPALA-7072: Impala does not support Heimdal Kerberos

None

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.

IMPALA-2422: % escaping does not work correctly when occurs at the end in a LIKE clause

If the final character in the RHS argument of a LIKE operator is an escaped \% character, it does not match a % final character of the LHS argument.

None

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

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

None

IMPALA-3094: Incorrect result due to constant evaluation in query with outer join

An OUTER JOIN query could omit some expected result rows due to a constant such as FALSE in another join clause. For example:

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

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-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-9350: Ranger audit logs for applying column masking policies missing

Impala is not producing these logs.

None

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 SQLGetDa ta call for column 1 returns NULL.

Fetch columns in the same order they are defined in the table.

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:

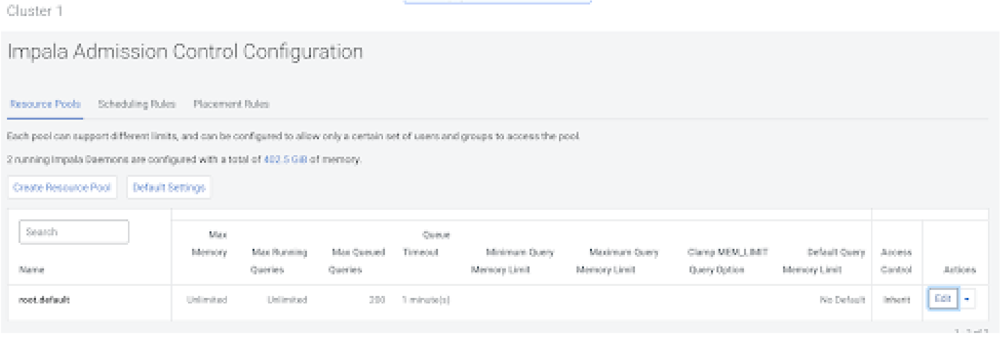
```
bin/start-impala-cluster.py
                              --impalad_args="--default_quer
y_options=fetch_rows_timeout_ms=0"
```

- 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:

  1. Click Clusters Impala Admission Control Configuration .

     You will see a screen similar to the following:

     Cluster 1

     Impala Admission Control Configuration

     Resource Pools | Scheduling Rules | Placement Rules

     Each pool can support different limits, and can be configured to allow only a certain set of users and groups to access the pool.

     2 running Impala Daemons are configured with a total of 402.5 GiB of memory.

     Create Resource Pool | Default Settings

     Search

Name	Max Memory	Max Running Queries	Max Queued Queries	Queue Timeout	Minimum Query Memory Limit	Maximum Query Memory Limit	Clamp MEM_LIMIT Query Option	Default Query Memory Limit	Access Control	Actions
root.default	Unlimited	Unlimited	200	1 minute(s)				No Default	Inherit	Edit

  2. Click Edit and scroll down to Default Query Options.
  3. Click + and add the query option FETCH_ROWS_TIMEOUT_MS = 0
- In CDW, users can configure their Impala VW by the following steps:

  1. Click the three dots and choose the Edit option.
  2. Select the Impala coordinator tab within the Configurations tab.
  3. Select flagfile in the drop-down menu.
  4. 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 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 need to override the bootstrap server URL by performing the following steps:

1. In Cloudera Manager, go to SMM Configuration Streams Messaging Manager Rest Admin Server Advanced Configuration Snippet (Safety Valve)
2. Override bootstrap server URL (hostname:port as set in the listeners for broker) for streams-messaging-manager.yaml.
3. Save your changes.
4. 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.

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 transacti
onal method because we are in an error state
              at org.apache.kafka.clients.producer.internals.Tr
ansactionManager.maybeFailWithError(TransactionManager.java:1125)
              at org.apache.kafka.clients.producer.internals.T
ransactionManager.maybeAddPartition(TransactionManager.java:442)
              at org.apache.kafka.clients.producer.KafkaProduce
r.doSend(KafkaProducer.java:1000)
              at org.apache.kafka.clients.producer.KafkaProduc
er.send(KafkaProducer.java:914)
```

```
              at org.apache.kafka.clients.producer.KafkaProducer
.send(KafkaProducer.java:800)
              .
              .
              .
              Caused by: org.apache.kafka.common.errors.Cluste
rAuthorizationException: 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.

CDPD-49304: AvroConverter does not support composite default values

AvroConverter cannot handle schemas containing a STRUCT type default value.

None.

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.

CFM-3532: The Stateless NiFi Source, Stateless NiFi Sink, and HDFS Stateless Sink connectors cannot use Snappy compression

This issue only affects Stateless NiFi Source and Sink connectors if the connector is running a dataflow that uses a processor that uses Hadoop libraries and is configured to use Snappy compression. The HDFS Stateless Sink connector is only affected if the Compression Codec or Compression Codec for Parquet properties are set to SNAPPY.

If you are affected by this issue, errors similar to the following will be present in the logs.

```
Failed to write to HDFS due to java.lang.UnsatisfiedLinkError: o
rg.apache.hadoop.util.NativeCodeLoader.buildSupportsSnappy()
```

```
Failed to write to HDFS due to java.lang.RuntimeException: nativ
e snappy library not available: this version of libhadoop was bu
ilt without snappy support.
```

Download and deploy missing libraries.

Important: Ensure that you complete steps 1-11 on all Kafka Connect hosts. Additionally, ensure that the advanced configuration snippet in step 12 is configured for all Kafka Connect role instances.

1. Create the /opt/nativelibs directory.

   ```
   mkdir /opt/nativelibs
   ```

2. Change the owner to kafka.

   ```
   chown kafka:kafka /opt/nativelibs
   ```

3. Locate the directory containing the Hadoop native libraries and copy its contents to the directory you created.

   ```
   cp /opt/cloudera/parcels/CDH/lib/hadoop/lib/native/* /opt/na
   tivelibs
   ```

4. Verify that libsnappy.so was copied to the directory you created.
5. Remove the following from /opt/nativelibs.

   ```
   libhadoop.a
                    libhadoop.so
                    libhadoop.so.1.0.0
   ```

6. Run the following command.

   ```
   hadoop version
   ```

   The command returns the Hadoop version running in the cluster. Note down the first three digits in the version.
7. Go to https://archive.apache.org/dist/hadoop/common/ and download the Hadoop version that matches the first three digits of the version running in the cluster.

   For example, if your Hadoop version is 3.1.1.7.1.9.0-296, then you need to download Hadoop 3.1.1.
8. Extract the downloaded archive.
9. Copy the following libraries from the downloaded archive to /opt/nativelibs on the cluster host.

   ```
   libhadoop.a
                    libhadoop.so.1.0.0
   ```

   The libraries are located in hadoop-*[***VERSION***]*/lib/native.
10. Create a symlink named libhadoop.so and point it to /opt/nativelibs/libhadoop.so.1.0.0.

    ```
    ln -s /opt/nativelibs/libhadoop.so.1.0.0 /opt/nativelibs/lib
    hadoop.so
    ```

11. Change the owner of every entry within /opt/nativelibs to kafka.

    ```
    chown -h kafka:kafka /opt/nativelibs/*
    ```

12. In Cloudera Manager, go to Kafka service Configuration .
13. Add the following key-value pair to Kafka Connect Environment Advanced Configuration Snippet (Safety Valve).

    - Key: LD_LIBRARY_PATH
    - Value: /opt/nativelibs
14. Click Save Changes.
15. Restart the Kafka service.

OPSAPS-69317: Kafka Connect Rolling Restart Check fails if SSL Client authentication is required

The rolling restart action does not work in Kafka Connect when the ssl.client.auth option is set to required. The health check fails with a timeout which blocks restarting the subsequent Kafka Connect instances.

You can set ssl.client.auth to requested instead of required and initiate a rolling restart again. Alternatively, you can perform the rolling restart manually by restarting the Kafka Connect instances one-by-one and checking periodically whether the service endpoint is available before starting the next one.

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 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_ent
ity_update_enabled=false
```

   Replace [KAFKA_SERVICE_NAME] with the service name of Kafka obtained in step 1. The service name should always be in lower case.

   c. Click Save Changes.

Known Issues in Kerberos

There are no known issues for Kerberos in Cloudera Runtime 7.3.1.0.

Known Issues in Apache Knox

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

CDPD-71305: Concurrent impala shell connection failure

If a user makes a concurrent impala-shell connection through Knox, then the connection fails.

Use only one Knox role.

CDPD-68146: Unable to update the log level for Knox from Cloudera Manager

Users are not able to change the log level for Knox from Cloudera Manager. Hence, it impacts debugging in case of any issue.

Change the level for the org.apache.knox.gateway logger in /var/lib/knox/gateway/conf/gateway-log4j2.xml file and restart Knox.

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-60379: During rolling upgrade of Knox service, access fails with 503/500/404/403 error code

The user operation which is performed during the rolling upgrade of knox might fail with 503/500/404/403 error code.

Retry the user operation.

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-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.

Known Issues in Apache Kudu

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

- Kudu HMS Sync is disabled and is not yet supported

Kudu supports both coarse-grain and fine-grain authorization, but Kudu does not yet support integration with Atlas.

None

KUDU-3619: Major delta compaction for a tablet might fail for particular workloads due to a bug introduced with KUDU-3367

A bug has been introduced with KUDU-3367 functionality. The bug manifests itself when a tablet server's maintenance thread attempts to run a major delta compaction on a tablet where many rows have been deleted, and the attempt fails with an error. To know more about the error message pattern, see KUDU-3619. If that happens, the corresponding tablet might accumulate a lot of updates that cannot be compacted and later garbage collected. In extreme cases, it could lead to running out of disk space when many tablet replicas hosted at the same tablet server hit the issue.

If a tablet server is affected by the issue, messages like the below are present in the tablet server's logs, where <tabletUUID> and <rowsetID> placeholders are populated with corresponding identifiers:

Major delta compaction failed on <tabletUUID>: Corruption: Failed major delta compaction on RowSet(<rowsetID>): No min key found: CFile base data in RowSet(<rowsetID>).

Set the --all_delete_op_delta_file_cnt_for_compaction flag to a very high value (e.g. 1000000) using the Tablet Server Advanced Configuration Snippet (Safety Valve) for gflagfile in the Cloudera Manager UI and restart all the tablet servers in the Kudu cluster.

Apache Jira: KUDU-3619

Known Issues in Apache Oozie

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

Oozie jobs fail (gracefully) on secure YARN clusters when JobHistory server is down

If the JobHistory server is down on a YARN (MRv2) cluster, Oozie attempts to submit a job, by default, three times. If the job fails, Oozie automatically puts the workflow in a SUSPEND state.

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

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

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

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

Unsupported Feature

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

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

Known Issues in Apache Parquet

There are no known issues for Parquet in Cloudera Runtime 7.3.1.0.

Known Issues in Apache Phoenix

There are no known issues for Phoenix in Cloudera Runtime 7.3.1.

Known Issues in Apache Ranger

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

CDPD-3296: Audit files for Ranger plugin components do not appear immediately in S3 after cluster creation

For Ranger plugin components (Atlas, Hive, HBase, etc.), audit data is updated when the applicable audit file is rolled over. The default Ranger audit rollover time is 24 hours, so audit data appears 24 hours after cluster creation.

To see the audit logs in S3 before the default rollover time of 24 hours, use the following steps to override the default value in the Cloudera Manager safety valve for the applicable service.

1. On the Configuration tab in the applicable service, select Advanced under CATEGORY.
2. Click the + icon for the <service_name> Advanced Configuration Snippet (Safety Valve) for ranger-<service_name>-audit.xml property.
3. Enter the following property in the Name box:

   xasecure.audit.destination.hdfs.file.rollover.sec.
4. Enter the desired rollover interval (in seconds) in the Value box. For example, if you specify 180, the audit log data is updated every 3 minutes.
5. Click Save Changes and restart the service.

Known Issues in Schema Registry

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

CDPD-40380: Authorization checking issue when Kerberos is disabled

Due to an issue in Ranger, when Kerberos is disabled then it is not possible to check authorization.

1. Open Schema Registry configuration in Cloudera Manager.
2. Find the ranger.plugin.schema-registry.service.name field.
3. Replace GENERATED_RANGER_SERVICE_NAME with the actual name of the service.
4. Restart the Schema Registry service.

CDPD-49304: AvroConverter does not support composite default values

AvroConverter cannot handle schemas containing a STRUCT type default value.

None.

OPSAPS-70971: Schema Registry does not have permissions to use Atlas after an upgrade

Following an upgrade, Schema Registry might not have the required permissions in Ranger to access Atlas. As a result, Schema Registry's integration with Atlas might not function in secure clusters where Ranger authorization is enabled.

1. Access the Ranger Console (Ranger Admin web UI).
2. Click the cm_atlas resource-based service.
3. Add the schemaregistry user to the all - * policies.
4. Click Manage Service Edit Service .
5. Add the schemaregistry user to the default.policy.users property.

OPSAPS-69317: Kafka Connect Rolling Restart Check fails if SSL Client authentication is required

The rolling restart action does not work in Kafka Connect when the ssl.client.auth option is set to required. The health check fails with a timeout which blocks restarting the subsequent Kafka Connect instances.

You can set ssl.client.auth to requested instead of required and initiate a rolling restart again. Alternatively, you can perform the rolling restart manually by restarting the Kafka Connect instances one-by-one and checking periodically whether the service endpoint is available before starting the next one.

Known Issues in Apache Solr

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

HBase indexer does not load netty and snappy libraries

The HBase indexer loads the netty and snappy libraries and these libraries are necessary for the Key-Value Indexer to work. However, the Key-Value Indexer cannot use these libraries if /tmp is mounted with the noexec property. To address this issue, you have to manually specify another directory instead of the default /tmp.

Perform the following steps to resolve this issue:

1. Go to the Key-Value Store Indexer service Configuration .
2. Search for the Key-Value Store Indexer Service Environment Advanced Configuration Snippet (Safety Valve) property.
3. If the HBASE_INDEXER_OPTS key is already present in the configuration, append the following value else add the following key and value:

```
Name: HBASE_INDEXER_OPTS
Value: -Dorg.apache.hbase.thirdparty.io.netty.native.workdir=/
var/hbase-solr/netty-workdir -Dorg.xerial.snappy.tempdir=/var/
hbase-solr/snappy-tempdir
```

Note:

- If the /var/hbase-solr/netty-workdir and /var/hbase-solr/snappy-tempdir file system locations do not exist, create the directories and ensure that the "hbase" user has permissions to write into these directories.
- Run the chown command on the directories. For example,

```
chown -R hbase:hbase /var/hbase-solr/netty-workdir
```

4. Restart the Key-Value Store Indexer service by clicking Key-Value Store Indexer service Actions Restart .

HBase Indexer does not work with JDK 17

Depending on the Cloudera Manager version used with CDP, HBase Indexer (KS Indexer) may have compatibility issues with JDK 17.

You have the following options to fix this issue:

- Upgrade Cloudera Manager to version 7.11.3 or higher.
- If upgrading Cloudera Manager is not an option, you can manually add the following to HBase Indexer Java options in Cloudera Manager:

```
--add-opens java.base/java.nio=ALL-UNNAMED --add-opens java.
base/java.util.concurrent.atomic=ALL-UNNAMED --add-opens jav
a.base/java.lang=ALL-UNNAMED --add-opens java.base/java.lang
.reflect=ALL-UNNAMED
```

Splitshard operation fails after CDH 6 to CDP upgrade

Collections are not reindexed during an upgrade from CDH 6 to CDP 7 because Lucene 8 (CDP) can read Lucene 7 (CDH 6) indexes.

If you try to execute a SPLITSHARD operation against such a collection, it fails with a similar error message:

```
o.a.s.h.a.SplitOp ERROR executing split: => java.lang.IllegalArg
umentException: Cannot merge a segment t
```

```
hat has been created with major version 7 into this index which
 has been created by major version 8
        at org.apache.lucene.index.IndexWriter.validateMergeRea
der(IndexWriter.java:3044)
java.lang.IllegalArgumentException: Cannot merge a segment that h
as been created with major version 7 into this index which has b
een created by major version 8
        at org.apache.lucene.index.IndexWriter.validateMergeReade
r(IndexWriter.java:3044) ~[lucene-core-8.11.2.7.1.9.3-2.jar:8.11
.2.7.1.9.3-2 a6ff93f9665115dffbdad0ad7f222fd1978d495d - jenkins -
 2023-12-02 00:05:23]
        at org.apache.lucene.index.IndexWriter.addIndexes(IndexWr
iter.java:3110) ~[lucene-core-8.11.2.7.1.9.3-2.jar:8.11.2.7.1.9.
3-2 a6ff93f9665115dffbdad0ad7f222fd1978d495d - jenkins - 2023-12
-02 00:05:23]
        at org.apache.solr.update.SolrIndexSplitter.doSplit(So
lrIndexSplitter.java:318) ~[solr-core-8.11.2.7.1.9.3-2.jar:8.11.
2.7.1.9.3-2 a6ff93f9665115dffbdad0ad7f222fd1978d495d - jenkins -
 2023-12-02 00:16:28]
        at org.apache.solr.update.SolrIndexSplitter.split(Solr
IndexSplitter.java:184) ~[solr-core-8.11.2.7.1.9.3-2.jar:8.11.2.
7.1.9.3-2 a6ff93f9665115dffbdad0ad7f222fd1978d495d - jenkins - 2
023-12-02 00:16:28]
        at org.apache.solr.update.DirectUpdateHandler2.split(Dir
ectUpdateHandler2.java:922) ~[solr-core-8.11.2.7.1.9.3-2.jar:8.1
1.2.7.1.9.3-2 a6ff93f9665115dffbdad0ad7f222fd1978d495d - jenkins
 - 2023-12-02 00:16:28]
        at org.apache.solr.handler.admin.SplitOp.execute(SplitOp
.java:165) ~[solr-core-8.11.2.7.1.9.3-2.jar:8.11.2.7.1.9.3-2 a6f
f93f9665115dffbdad0ad7f222fd1978d495d - jenkins - 2023-12-02 00:
16:28]
        at org.apache.solr.handler.admin.CoreAdminOperation.execu
te(CoreAdminOperation.java:367) ~[solr-core-8.11.2.7.1.9.3-2.jar
:8.11.2.7.1.9.3-2 a6ff93f9665115dffbdad0ad7f222fd1978d495d - jen
kins - 2023-12-02 00:16:28]
```

This happens because the segment created using a Lucene 7 index cannot be merged into a Lucene 8 index.

Drop the entire collection, delete the data in HDFS and recreate the collection with Solr 8 configs.

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.MasterRegistryFetc
hException: Exception making rpc to masters [quasar-bmyccr-2.qua
sar-bmyccr.root.hwx.site,22001,-1]
        at org.apache.hadoop.hbase.client.MasterRegistry.lambda$g
roupCall$1(MasterRegistry.java:244)
        at org.apache.hadoop.hbase.util.FutureUtils.lambda$addLi
stener$0(FutureUtils.java:68)
        at java.util.concurrent.CompletableFuture.uniWhenCompl
ete(CompletableFuture.java:774)
        at java.util.concurrent.CompletableFuture.uniWhenComplet
eStage(CompletableFuture.java:792)
        at java.util.concurrent.CompletableFuture.whenComplete(Co
mpletableFuture.java:2153)
        at org.apache.hadoop.hbase.util.FutureUtils.addListener(F
utureUtils.java:61)
```

```
        at org.apache.hadoop.hbase.client.MasterRegistry.groupCa
ll(MasterRegistry.java:228)
        at org.apache.hadoop.hbase.client.MasterRegistry.call(Ma
sterRegistry.java:265)
        at org.apache.hadoop.hbase.client.MasterRegistry.getMetaR
egionLocations(MasterRegistry.java:282)
        at org.apache.hadoop.hbase.client.ConnectionImplementati
on.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.

Unable to see single valued and multivalued empty string values when querying collections after upgrade to CDP

After upgrading from CDH or HDP to CDP, you are not able to see single valued and multi Valued empty string values in CDP.

This behavior in CDP is due to the remove-blank processor present in solrconfig.xml in Solr 8.

Remove the remove-blank processor from solrconfig.xml.

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.tempdir 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.tempdir=/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
```

Mergeindex operation with --go-live fails after CDH 6 to CDP upgrade

During an upgrade from CDH6 to CDP, collections are not reindexed because Lucene 8 (CDP) can read Lucene 7 (CDH6) indexes.

If you try to execute MapReduceIndexerTool (MRIT) or HBase Indexer MRIT with --go-live against such a collection, you get a similar error message:

```
Caused by: java.lang.IllegalArgumentException: Cannot merge a se
gment that has been created with major version 8 into this index
 which has been created by major version 7
        at org.apache.lucene.index.IndexWriter.validateMergeReade
r(IndexWriter.java:2894)
        at org.apache.lucene.index.IndexWriter.addIndexes(Index
Writer.java:2960)
        at org.apache.solr.update.DirectUpdateHandler2.mergeIn
dexes(DirectUpdateHandler2.java:570)
        at org.apache.solr.update.processor.RunUpdateProcessor.
processMergeIndexes(RunUpdateProcessorFactory.java:95)
        at org.apache.solr.update.processor.UpdateRequestProcesso
r.processMergeIndexes(UpdateRequestProcessor.java:63)
```

This happens because CDP MRIT and HBase indexer use Solr 8 as embedded Solr, which creates a Lucene 8 index. It cannot be merged (using MERGEINDEXES) into an older Lucene 7 index.

In the case of MRIT the only way to move past this issue is to drop the entire collection, delete the data in HDFS and recreate the collection with Solr 8 configs.

For HBase Indexer MRIT an alternative workaround is setting the number of reducers to 0 (--reducers 0) because in this case documents are sent directly from the mapper tasks to live Solr servers instead of using MERGEINDEXES.

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-28006: Solr access via Knox fails with impersonation error though auth_to_local and proxy user configs are set

Currently the names of system users which are impersonating users with Solr should match with the names of their respective Kerberos principals.

If, for some reason, this is not feasible, you must add the user name you want to associate with the custom Kerberos principal to Solr configuration via the Solr Service Environment Advanced Configuration Snippet (Safety Valve) environment variable in Cloudera Manager.

For more information, see Configuring custom Kerberos principals and custom system users.

CDH-77598: Indexing fails with socketTimeout

Starting from CDH 6.0, the HTTP client library used by Solr has a default socket timeout of 10 minutes. Because of this, if a single request sent from an indexer executor to Solr takes more than 10 minutes to be serviced, the indexing process fails with a timeout error.

This timeout has been raised to 24 hours. Nevertheless, there still may be use cases where even this extended timeout period proves insufficient.

If your MapreduceIndexerTool or HBaseMapreduceIndexerTool batch indexing jobs fail with a timeout error during the go-live (Live merge, MERGEINDEXES) phase (This means the merge takes longer than 24 hours).

Use the --go-live-timeout option where the timeout can be specified in milliseconds.

CDPD-12450: CrunchIndexerTool Indexing fails with socketTimeout

The http client library uses a socket timeout of 10 minutes. The Spark Crunch Indexer does not override this value, and in case a single batch takes more than 10 minutes, the entire indexing job fails. This can happen especially if the morphlines contain DeleteByQuery requests.

Try the following workarounds:

- Check the batch size of your indexing job. Sending too large batches to Solr might increase the time needed on the Solr server to process the incoming batch.
- If your indexing job uses deleteByQuery requests, consider using deleteById wherever possible as deleteByQuery involves a complex locking mechanism on the Solr side which makes processing the requests slower.
- Check the number of executors for your Spark Crunch Indexer job. Too many executors can overload the Solr service. You can configure the number of executors by using the --mappers parameter
- Check that your Solr installation is correctly sized to accommodate the indexing load, making sure that the number of Solr servers and the number of shards in your target collection are adequate.
- The socket timeout for the connection can be configured in the morphline file. Add the solrClientSocketTimeout parameter to the solrLocator command

  Example

  ```
  SOLR_LOCATOR :
  {
    collection : test_collection
    zkHost : "zookeeper1.example.corp:2181/solr"
  # 10 minutes in milliseconds
    solrClientSocketTimeout: 600000
    # Max number of documents to pass per RPC from morphline to
   Solr Server
    # batchSize : 10000
  }
  ```

CDPD-29289: HBaseMapReduceIndexerTool fails with socketTimeout

The http client library uses a socket timeout of 10 minutes. The HBase Indexer does not override this value, and in case a single batch takes more than 10 minutes, the entire indexing job fails.

You can overwrite the default 600000 millisecond (10 minute) socket timeout in HBase indexer using the --solr-client-socket-timeout optional argument for the direct writing mode (when the value of the --reducers optional argument is set to 0 and mappers directly send the data to the live Solr).

CDPD-20577: Splitshard operation on HDFS index checks local filesystem and fails

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

Run the following command to skip the check for sufficient disk space altogether:

- On nonsecure clusters:

```
curl 'http://$[***SOLR_SERVER_HOSTNAME***]:8983/so
lr/admin/collections?action=SPLITSHARD&collectio
n=[***COLLECTION_NAME***]&shard=[***SHARD_TO_SPLIT***]&skipFre
eSpaceCheck=true'
```

- On secure clusters:

```
curl -k -u : --negotiate 'http://
$[***SOLR_SERVER_HOSTNAME***]:8985/solr/admin/collections
?action=SPLITSHARD&collection=[***COLLECTION_NAME***]&sha
rd=[***SHARD_TO_SPLIT***]&skipFreeSpaceCheck=true'
```

Replace *[***SOLR_SERVER_HOSTNAME***]* with a valid Solr server hostname, *[***COLLECTION_NAME***]* with the collection name, and *[***SHARD_TO_SPLIT***]* with the ID of the to split.

To verify that the command executed succesfully, check overseer logs for a similar entry:

```
2021-02-02 12:43:23.743 INFO  (OverseerThreadFactory-9-thread-5-
processing-n:myhost.example.com:8983_solr) [c:example s:shard1
  ] o.a.s.c.a.c.SplitShardCmd Skipping check for sufficient disk
 space
```

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.

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.

CDPD-13923: Every Configset is Untrusted Without Kerberos

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

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

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

CDPD-71422: Solr went into an unhealthy state after the data lake upgrade

After the Data Lake upgrade to the 7.3.1.0 version, the Solr service becomes unhealthy for the public cloud environments (AWS, Azure, and GCP). This is an intermittent issue.

Manually restart the Solr service in the Data Lake after an upgrade.

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.

Spark 3: RAPIDS Accelerator is not available

The RAPIDS Accelerator for Apache Spark is currently not available in Cloudera Public Cloud 7.3.1.0.

Workaround: None

The CHAR(n) type handled inconsistently, depending on whether the table is partitioned or not.

In upstream Spark 3 the spark.sql.legacy.charVarcharAsString configuration was introduced, but it does not solve all incompatibilities with Spark 2.

Workaround: None. A new configuration spark.cloudera.legacy.charVarcharLegacyPadding will be introduced in a future version to keep compatibility with Spark 2, but it isn't available in 7.3.1.

Note: The CHAR type is legacy in SQL, and using it is discouraged. Cloudera recommends using VARCHAR or STRING instead.

Apache Jira: SPARK-33480

Known Issues for Apache Sqoop

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

CDPD-44431: Using direct mode causes problems

Using direct mode has several drawbacks:

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

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

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

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

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

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

Apache JIRA: None

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

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

PARQUET-99

Known issues in Streams Messaging Manager

Learn about the known issues in Streams Messaging Manager (SMM), the impact or changes to the functionality, and the workaround.

OPSAPS-59597: SMM UI logs are not supported by Cloudera Manager

Cloudera Manager does not display a Log Files menu for SMM UI role (and SMM UI logs cannot be displayed in the Cloudera Manager UI) because the logging type used by SMM UI is not supported by Cloudera Manager.

View the SMM UI logs on the host.

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.

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 need 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:

1. In Cloudera Manager, go to SMMConfigurationStreams Messaging Manager Rest Admin Server Advanced Configuration Snippet (Safety Valve) for streams-messaging-manager.yaml.
2. Add the following value for bootstrap servers.

   ```
   streams.messaging.manager.kafka.bootstrap.servers=<comma-sep
   arated list of brokers>
   ```
3. Save your changes.
4. Restart SMM.

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.

Limitations

CDPD-36422: 1MB flow.snapshot freezes Safari

While importing large connector configurations, flow.snapshots reduces the usability of the Streams Messaging Manager when using Safari browser.

Use a different browser (Chrome/Firefox/Edge).

Known Issues in Streams Replication Manager

Learn about the known issues in Streams Replication Manager (SRM), 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, it's 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 [TARGE
T_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.

SRM cannot ensure the exactly-once semantics of transactional source topics

SRM data replication uses at-least-once guarantees, and as a result cannot ensure the exactly-once semantics (EOS) of transactional topics in the backup/target cluster.

Note: Even though EOS is not guaranteed, you can still replicate the data of a transactional source, but you must set isolation.level to read_committed for SRM's internal consumers. This can be done by adding _[***SOURCE CLUSTER ALIAS***]->[***TARGET CLUSTER ALIAS***]_.consumer.isolation.level=read_committed to the Streams Replication Manager's Replication Configs SRM service property in Cloudera Manger. The isolation.level property can be set on a global connector or replication level. For example:

```
#Global connector level
                  connectors.consumer.isolation.leve
l=read_committed
                  #Replication level
                  uswest->useast.consumer.isolation.le
vel=read_committed
```

SRM checkpointing is not supported for transactional source topics

SRM does not correctly translate checkpoints (committed consumer group offsets) for transactional topics. Checkpointing assumes that the offset mapping function is always increasing, but with transactional source topics this is violated. Transactional topics have control messages in them, which take up an offset in the log, but they are never returned on the consumer API. This causes the mappings to decrease, causing issues in the checkpointing feature. As a result of this limitation, failover operations for transactional topics is not possible.

Known Issues in YARN, YARN Queue Manager and MapReduce

Learn about the known issues in YARN, YARN Queue Manager and MapReduce, 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:cdrwa which is less secure.

Third party applications do not launch if MapReduce framework path is not included in the client configuration

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

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

JobHistory URL mismatch after server relocation

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

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

CDH-6808: Routable IP address required by ResourceManager

ResourceManager requires routable host:port addresses for yarn.resourcemanager.scheduler.address, and does not support using the wildcard 0.0.0.0 address.

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

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

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

To view the history for a killed Spark application, see the Spark HistoryServer web UI instead.

COMPX-3329: Autorestart is not enabled for Queue Manager in Data Hub

In a Data Hub cluster, Queue Manager is installed with autorestart disabled. Hence, if Queue Manager goes down, it will not restart automatically.

If Queue Manager goes down in a Data Hub cluster, you must go to the Cloudera Manager Dashboard and restart the Queue Manager service.

COMPX-5817: Queue Manager UI will not be able to present a view of pre-upgrade queue structure. CM Store is not supported and therefore Yarn will not have any of the pre-upgrade queue structure preserved.

When a Data Hub cluster is deleted, all saved configurations are also deleted. All YARN configurations are saved in CM Store and this is yet to be supported in Data Hub and Cloudera Manager. Hence, the YARN queue structure also will be lost when a Data Hub cluster is deleted or upgraded or restored.

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.

CDPD-75652: Reverse DNS lookup fails for YARN but works for HDFS

Submitting a YARN application from a host without proper DNS setup (reverse DNS does not work for the YARN ResourceManager's host) results in the Server has invalid Kerberos principal error.

Add the following to YARN Service Advanced Configuration Snippet (Safety Valve) for the yarn-site.xml file:

```
<property>
<name>yarn.resourcemanager.principal.pattern</name>
<value>*</value>
</property>
```

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 ZooKeeper

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

Zookeeper-client does not use ZooKeeper TLS/SSL automatically

The command-line tool ‘zookeeper-client’ is installed to all Cloudera Nodes and it can be used to start the default Java command line ZooKeeper client. However even when ZooKeeper TLS/SSL is enabled, the zookeeper-client command connects to localhost:2181, without using TLS/SSL.

Workaround:

Manually configure the 2182 port, when zookeeper-client connects to a ZooKeeper cluster.The following is an example of connecting to a specific three-node ZooKeeper cluster using TLS/SSL:

```
CLIENT_JVMFLAGS="-Dzookeeper.clientCnxnSocket=org.apache.zoo
keeper.ClientCnxnSocketNetty -Dzookeeper.ssl.keyStore.locati
on=<path to your configured keystore> -Dzookeeper.ssl.keyStor
e.password=<the password you configured for the keystore>  -
Dzookeeper.ssl.trustStore.location=<path to your configured
 truststore> -Dzookeeper.ssl.trustStore.password=<the password
 you configured for the truststore> -Dzookeeper.client.secu
re=true" zookeeper-client -server <your.zookeeper.server-1>:218
2,<your.zookeeper.server-2>:2182,<your.zookeeper.server-3>:2182
```

Behavioral Changes In Cloudera Runtime 7.3.1

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

Behavioral Changes in Atlas

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

Summary:

The Exclude SubTypes and Exclude Sub-classifications filters were removed from the **Table** tab of entity details.

Previous behavior:

Previously, the Exclude SubTypes and Exclude Sub-classifications filters were available from the **Table** tab in entity details. There were no properties being passed to these filters when you visited the entity details of the page.

New behavior:

The two unused filter checkboxes Exclude SubTypes and Exclude Sub-classifications from the **Table** tab of entity detail page were removed.

Summary:

Special character validation was added to glossary, term and category names in Apache Atlas.

Previous behavior:

The special characters ('@', '.', '<', '>') could be used in glossary, term and category name fields.

New behavior:

The special characters ('@', '.', '<', '>') are no longer accepted in glossary, term and category name fields by the validation introduced. Avoid using these characters when creating glossary names, glossary terms and category names.

Behavioral Changes in Hive

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

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.compactor.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 Impala

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

Summary:

Impala now unregisters timed-out queries promptly to free memory, retaining error messages for clients that return later.

Previous behavior:

Timed-out queries remained registered until the session closed, keeping memory occupied and sometimes leaving failed queries in an active state if not explicitly closed.

New behavior:

Timed-out queries are unregistered immediately to free memory, while error messages are kept in a new structure so clients can still receive an error message if they return later.

Apache Jira: IMPALA-12602

Behavioral Changes in Knox

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

Knox token impersonation config

Summary

Knox token service has been changed to use the identity assertion provider configuration for impersonation.

Previous behaviour

The token service had its own impersonation configuration.

New behaviour

The token service relies on the identity assertion provider for impersonation configuration.

PEM file name change

Summary

The name of the pem file generated through knoxcli.sh has been changed.

Previous behaviour

The name of the file was gateway-identity.pem.

New behaviour

The name of the file is now gateway-client-trust.pem.

Composite authorization provider misconfiguration

Summary

Composite authorization provider misconfiguration behavior

Previous behaviour

1. If composite.provider.names is empty, the topology would fail deployment.
2. If composite.provider.names has an invalid value, the topology would fail deployment.

New behaviour

1. Deployment succeeds, and Knox allows access with no authorization since none is configured.
2. Deployment succeeds, but Knox rejects requests with a HTTP 403 response because the configuration is present (indicating that authorization is expected) but invalid.

Inactive topologies

Summary

Knox distinguishes inactive topologies from undeployed topologies.

Previous behaviour

Requests for topologies which are not yet fully deployed result in HTTP 404 responses.

New behaviour

Requests for topologies which are not yet fully deployed result in HTTP 503 responses.

Behavioral Changes in Livy

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

Summary:

The Livy proxy user is taken from Livy for Spark 3's configuration.

Previous behavior:

The custom Kerberos principal configuration was updated via the Livy service.

New behavior:

The Livy proxy user is taken from Livy for Spark 3's configuration, as the Livy service has been replaced with Livy for Spark3 in Cloudera Public Cloud version 7.3.1.

Behavioral Changes in Ranger

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

Summary: Ranger access audit behavior changes.

Previous behavior:

When you ran `hdfs dfs -copyFromLocal` command, audit logs were generated for the following:

- "write" Access Type and "write" permission.
- "rename" Access Type and "write" permission.
- "rename" Access Type and "write" permission.

When you ran `hdfs dfs -touch` command, audit log was generated for the following:

- "write" Access Type and "write" permission.

New behavior:

When you run `hdfs dfs -copyFromLocal` command, audit logs are generated for the following:

- "create" Access Type and "write" permission.
- "rename" Access Type and "write" permission.

When you run `hdfs dfs -touch` command, audit log is generated for the following:

- "create" Access Type and "write" permission.

Behavioral Changes in Spark

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

Summary:

Spark 2 has been removed from Cloudera Runtime.

Previous behavior:

Spark 2 was the default version in Cloudera Runtime, Spark 3 was available as an add-on parcel.

New behavior:

Spark 3 is the default Spark version in Cloudera Runtime. Spark 2 has been removed and no longer available in 7.3.1.0.

⚠ **Important:**

Spark 3 contains a large number of changes from Spark 2.

Refer to *Upgrading Spark* for more information on upgrading Spark clusters to 7.3.1.0, and *Migrating Spark Applications* for more information on migrating your existing Spark applications between versions 2 and 3.

Related Information

Upgrading Spark

Migrating Spark Applications

Deprecation Notices In Cloudera Runtime 7.3.1

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

Terminology

Items in this section are designated as follows:

Deprecated

Technology that Cloudera is removing in a future CDP release. Marking an item as deprecated gives you time to plan for removal in a future CDP release.

Moving

Technology that Cloudera is moving from a future CDP release and is making available through an alternative Cloudera offering or subscription. Marking an item as moving gives you time to plan for removal in a future CDP release and plan for the alternative Cloudera offering or subscription for the technology.

Removed

Technology that Cloudera has removed from CDP and is no longer available or supported as of this release. Take note of technology marked as removed since it can potentially affect your upgrade plans.

Removed Components and Product Capabilities

- Apache Spark 2

  Spark 3 is the default Spark version in Cloudera Runtime. Spark 2 (and Livy 2) has been removed and no longer available in 7.3.1

  ⚠ **Important:**

  Spark 3 contains a large number of changes from Spark 2.

  Refer to *Upgrading Spark* for more information on upgrading Spark clusters to 7.3.1.0, and *Migrating Spark Applications* for more information on migrating your existing Spark applications between versions 2 and 3.
- Apache Livy 2 (see Deprecation Notices for Apache Livy)
- Apache Zeppelin (see Deprecation Notices for Apache Zeppelin)

Please contact Cloudera Support or your Cloudera Account Team if you have any questions.

Related Information

Upgrading Spark

Migrating Spark Applications

Platform and OS

The listed Operating Systems, databases, and instant client library are deprecated or removed from the 7.3.1 release.

Database Support:

The listed databases are deprecated from the 7.3.1 release:

- None

The following database is removed and no longer supported from the 7.3.1 release:

- PostgreSQL 11

Operating System

The listed operating systems are deprecated from the 7.3.1 release:

- None

The following operating system is removed and no longer supported from the 7.3.1 release:

- CentOS

Note: CentOS Linux 7 has reached end of life. Ensure to migrate to RHEL/Oracle Linux or any supported operating system before upgrading to 7.3.1.

Deprecation Notices for Apache Kafka

Certain features and functionality in Apache Kafka are deprecated or removed in Cloudera Runtime 7.3.1. 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 Livy

Certain features and functionality in Apache Livy are deprecated or removed in Cloudera Runtime 7.3.1. You must review these changes along with the information about the features in Livy that will be removed or deprecated in a future release.

Removed

Apache Livy 2

As Spark 3 is the default Spark version in Cloudera Runtime, Livy 2 has been removed, alongside with Spark 2, and no longer available in 7.3.1

Important:

Spark 3 contains a large number of changes from Spark 2.

For more information on upgrading to Spark 3, refer to Upgrading Spark for more information on upgrading Spark clusters to 7.3.1, and Migrating Spark Applications for more information on migrating your existing Spark applications between versions 2 and 3.

Deprecation Notices for Apache Oozie

Certain features and functionality in Apache Oozie are deprecated or removed in Cloudera Runtime 7.3.1. 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 7.3.1, Cloudera decided to deprecate Oozie Spark actions, which are based on Spark 2. Consequently, Oozie's Spark actions will no longer be available, and if you attempt to execute a Spark action, an error will be raised.

Starting from 7.3.1, you must migrate to Spark 3 to use Spark actions. For more information, see Spark 3 support in Oozie.

Deprecation Notices for Apache Spark

Certain features and functionality in Apache Spark 2 are deprecated or removed in Cloudera Runtime 7.3.1. You must review these changes along with the information about the features in Spark 2 that will be removed or deprecated in a future release.

Removed

Apache Spark 2

Spark 3 is the default Spark version in Cloudera Runtime. Spark 2 has been removed and no longer available in 7.3.1

Important:

Spark 3 contains a large number of changes from Spark 2.

Refer to Upgrading Spark for more information on upgrading Spark clusters to 7.3.1, and Migrating Spark Applications for more information on migrating your existing Spark applications between versions 2 and 3.

Deprecation Notices for Apache Zeppelin

Certain features and functionality in Apache Zeppelin are deprecated or removed in Cloudera Runtime 7.3.1. You must review these changes along with the information about the features in Zeppelin that will be removed or deprecated in a future release.

Removed

Apache Zeppelin

Apache Zeppelin is removed from Cloudera Public Cloud.

Cloudera recommends you back up all existing Zeppelin notebooks before upgrading to version 7.3.1.

Fixed Common Vulnerabilities and Exposures 7.3.1

Common vulnerabilities and Exposures (CVEs) fixed in this release.

- CVE-2023-6378 Logback
- CVE-2023-6481 Logback
- CVE-2023-2976 Google Guava
- CVE-2020-8908 Google Guava
- CVE-2018-10237 Google Guava
- CVE-2023-52428 Nimbus-jose-jwt
- CVE-2023-45865 Akka-actor

- CVE-2021-42697 Akka-http-core
- CVE-2021-23339 Akka-http-core
- CVE-2022-31023 Akka-http-server
- CVE-2021-26291 Apache Maven
- CVE-2022-46337 Apache Derby
- CVE-2023-22006 Graal-sdk
- CVE-2023-50386 Apache Solr
- CVE-2023-50291 Apache Solr
- CVE-2023-50292 Apache Solr
- CVE-2023-50298 Apache Solr
- CVE-2023-1932 Hibernate Validator
- CVE-2024-22201 Eclipse Jetty
- CVE-2024-21634 Amazon Ion
- CVE-2017-7525 Jackson-mapper-asl
- CVE-2019-10172 Jackson-mapper-asl
- CVE-2023-51775 Jose4j
- CVE-2020-15522 Bouncycastle
- CVE-2020-0187 Bouncycastle
- CVE-2022-1471 Snakeyaml
- CVE-2022-25857 Snakeyaml
- CVE-2022-38749 Snakeyaml
- CVE-2022-38751 Snakeyaml
- CVE-2022-38752 Snakeyaml
- CVE-2022-41854 Snakeyaml
- CVE-2022-38750 Snakeyaml
- CVE-2021-31684 Json-smart
- CVE-2023-1370 Json-smart
- CVE-2021-27568 Json-smart
- CVE-2021-4178 Fabric 8 Kubernetes client
- CVE-2023-3635 Okio
- CVE-2024-1597 Postgresql
- CVE-2023-45857 Axios
- CVE-2022-4244 Plexus-utils
- CVE-2022-4245 Plexus-utils
- CVE-2023-34453 Snappy-java
- CVE-2023-34454 Snappy-java
- CVE-2023-34455 Snappy-java
- CVE-2023-43642 Snappy-java
- CVE-2023-34042 Spring Security
- CVE-2024-22257 Spring Security
- CVE-2023-20859 Spring Vault
- CVE-2024-22243 Spring Framework
- CVE-2024-22262 Spring Framework
- CVE-2024-22259 Spring Framework
- CVE-2024-1300 Vertx-core
- CVE-2023-44483 Xmlsec
- CVE-2024-31573 Xmlunit-core
- CVE-2024-38998 Requirejs
- CVE-2024-38999 Requirejs
- CVE-2023-4759 Eclipse Jgit