Cloudera ODBC Connector for Apache Spark 2.6.21

Released 2021-12-17

These release notes provide details of enhancements, features, known issues, and workflow changes in Cloudera ODBC Connector for Apache Spark 2.6.21, as well as the version history.

For information about upcoming support deprecations or removals, see the Workflow Changes section. Deprecated features will not receive any updates, but will continue to be usable in their current state until support is removed in a future release.

Enhancements & New Features

[SPARKO-896] Support for CDP 7.1

The connector now supports CDP 7.1. Note that Livy Thrift Server must be running in the cluster for the connector to function correctly.

Known Issues

The following are known issues that you may encounter due to limitations in the data source, the connector, or an application.

- [SPARKO-879] When connecting to a server that supports multiple catalogs, the connector no longer reports the catalog for schemas and tables as SPARK.

  The Spark server now reports the catalog.

- [SPARKO-670] In some cases, when retrieving timestamp data, the connector returns an error.

  In some cases, when connecting to certain distributions of Apache Spark, the connector returns the following error: "Conversion from number to string failed due to undersized character buffer". This issue affects versions 2.6.12 to 2.6.14 of the Spark ODBC connector.

  As a workaround, set `EnableArrow=0` in the connection string or DSN.

- [SPARKO-620] Issue with date and timestamp before the beginning of the Gregorian calendar when connecting to Spark 2.4.4 or later, or versions previous to 3.0, with Arrow result set serialization.

When using Spark 2.4.4 or later, or versions previous to Spark 3.0, DATE and TIMESTAMP data before October 15, 1582 may be returned incorrectly if the server supports serializing query results using Apache Arrow. This issue should not impact most distributions of Apache Spark.

To confirm if your distribution of Spark 2.4.4 or later has been impacted by this issue, you can execute the following query:

SELECT DATE '1581-10-14'

If the result returned by the connector is 1581-10-24, then you are impacted by the issue. In this case, if your data set contains date and/or timestamp data earlier than October 15, 1582, you can work around this issue by adding EnableArrow=0 in your DSN or connection string to disable the Arrow result set serialization feature.

- When retrieving data from a BINARY column, a ClassCastException error occurs.

  In Spark 1.6.3 or earlier, the server sometimes returns a ClassCastException error when attempting to retrieve data from a BINARY column.

  This issue is resolved as of Spark 2.0.0.

  For more information, see the JIRA issue posted by Apache named "When column type is binary, select occurs ClassCastException in Beeline" at https://issues.apache.org/jira/browse/SPARK-12143.

Workflow Changes

The following changes may disrupt workflows from earlier versions.

In addition to changes that are already implemented in the current version of the connector, this section describes potentially disruptive changes that will be implemented in a future version of the connector, so that you can plan accordingly.

Upcoming

[SPARKO-585][SPARKO-587] Removing support for Spark 1.6, 2.1, and 2.2

As early as August 2021, the connector will no longer support servers that run Spark version 1.6, 2.1, or 2.2. For information about the supported Spark versions, see the *Installation and Configuration Guide*.

Contact Us

If you are having difficulties using the driver, our Community Forum may have your solution. In addition to providing user to user support, our forums are a great place to share your questions, comments, and feature requests with us.

If you are a Subscription customer you may also use the Cloudera Support Portal to search the Knowledge Base or file a Case.

Important: To help us assist you, prior to contacting Cloudera Support please prepare a detailed summary of the client and server environment including operating system version, patch level, and configuration.