

CDP One

CDP One Release Notes

Date published: 2022-06-03

Date modified: 2022-08-15

CLOUDERA

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

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

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

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

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

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

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

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

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

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

Contents

CDP One Known Issues and Limitations..... 4

CDP One Known Issues and Limitations

This section lists the known issues and limitations applicable to CDP One.

HMS Partition Discovery is disabled for both DL and DH

Issue:

HMS Partition Discovery is disabled for both data lakes and data hubs.

Workaround

Run the MSCK (metastore consistency check) Hive command every time you need to synchronize a partition with the file system. For more information see [Partitioned tables](#).

Spark connection via SparklyR can only read external hive tables

Issue:

Spark connections via SparklyR can only read external hive tables (not managed ACID hive tables).

Workaround

To access Hive from Spark, you need to use the Hive Warehouse Connector (HWC) implicitly or explicitly.

Spark and Hive tables interoperate using the Hive Warehouse Connector and Spark Direct Reader to access ACID managed tables. The Hive Warehouse Connector is designed to access managed ACID v2 Hive tables from Spark. HWC is a Spark library/plugin that is launched with the Spark app.

You can access external tables from Spark directly using SparkSQL. You do not need HWC to read or write Hive external tables. Spark users just read from or write to Hive directly. You can read Hive external tables in ORC or Parquet formats. You can write Hive external tables in ORC format only.

Use the Spark Direct Reader and HWC for ETL jobs. For other jobs, consider using Apache Ranger and the HiveWarehouseConnector library to provide row and column, fine-grained access to the data.

HWC supports spark-submit and pyspark. The spark thrift server is not supported.

To deploy JAR files for Sqoop, contact Cloudera Support

Issue:

Errors may occur if you attempt to add JAR files to the Sqoop directory.

Workaround

Contact Cloudera support for help with deploying JAR files for Sqoop.