

Cloudera Runtime 7.2.18

Hive Troubleshooting

Date published: 2022-12-16

Date modified: 2024-03-20

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

Unable to alter S3-backed tables.....	4
--	----------

Unable to alter S3-backed tables

Condition

In CDP Public Cloud on AWS or wherever S3-backed tables are used, an ALTER TABLE <table-name> RENAME TO <new-table-name> operation might result in an error stating that the new table location already exists.

Cause

A rename operation on an S3-backed table is not instant. It has to move each file in the table to a new location one at a time. Depending on how many files are in the table directory, this could take more than 5 minutes, which is the default timeout specified.

If the time specified for the connection elapses before the rename completes, Hive tries again and fails because the new directory already exists (from the previous attempt).

Because of this retry behavior, an end-user job or query can result in a misleading error:

```
Caused by: java.sql.SQLException: Error while compiling statement: FAILED: Execution Error, return code 1
from org.apache.hadoop.hive.ql.ddl.DDLTask. Unable to alter table. New location for this table
hive.<dbname>.<tablename> already exists : s3a://<bucket>/warehouse/tablespace/managed/hive/<dbname>.db/<tablename>
```


Solution

Procedure

Check the Hive Metastore logs for the time period close to the failed rename and look for the following timeout message:

```
ERROR org.apache.hadoop.hive.metastore.RetryingHMSHandler: [pool-7-thread-31549]: Error happens in method alter_table_req:
MetaException(message:Timeout when executing method: alter_table_req; 694779ms exceeds 600000ms)
```

If there is a timeout message, perform the following steps to increase the value of the timeout property for both Hive clients and Hive Metastore:

1. Log in to Cloudera Manager as an administrator.
2. Go to Clusters Hive Configuration and search for the 'Hive Metastore Connection Timeout' property.
3. Modify the value to 1 hour and click Save Changes.
4. Go to Clusters Hive Metastore Configuration and search for 'Hive Metastore Server Advanced Configuration Snippet (Safety Valve) for hive-site.xml'.
5. Click  and add the following property:

```
Name: hive.metastore.client.socket.timeout
Value: 3600
```

6. Click Save Changes.
7. Restart the Hive and Hive Metastore roles.