

# Troubleshooting Cloudera Data Catalog

Date published: 2019-11-14

Date modified: 2025-10-17



# Legal Notice

© Cloudera Inc. 2025. 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

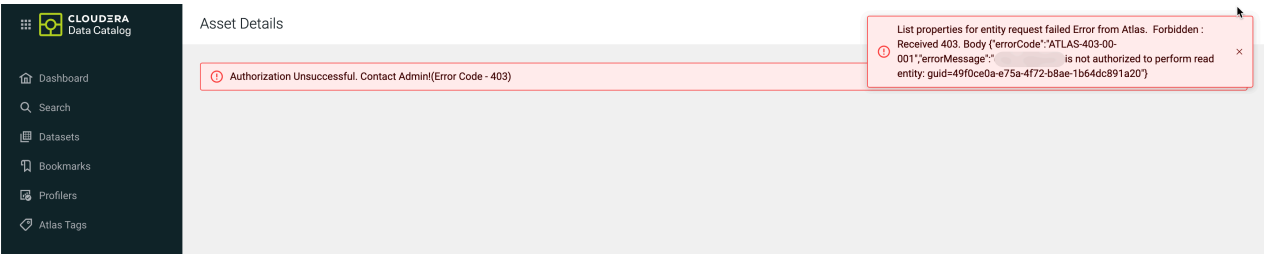
|  |          |
|--|----------|
| <b>Missing authorization for viewing assets.....</b> | <b>4</b> |
| <b>FAQs.....</b>                                     | <b>6</b> |
| <b>Associating tags to Hive assets.....</b>          | <b>6</b> |

# Missing authorization for viewing assets

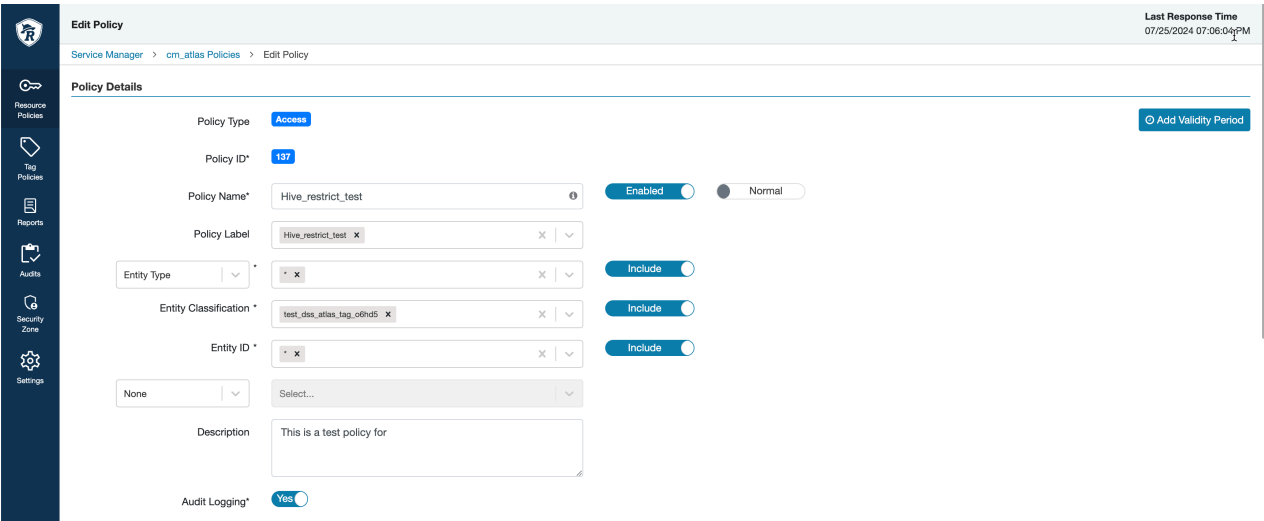
Cloudera Data Catalog users must have appropriate authorization set up in Apache Ranger to view certain asset types.

## Lack of permission to view a dataset

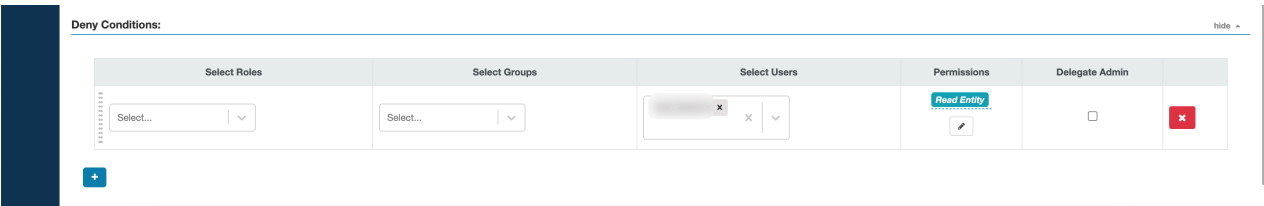
If you do not have access to a dataset, a similar error message to the following can be seen:



The previous error message is the result of a Hive policy setting similar to the following:



The following image displays the **Deny Conditions** set for the specific user.



## Hive Ranger policy

You must set up Hive Ranger resource-based policies as per your work requirements for Hive assets in Cloudera Data Catalog.

The following diagram provides a sample Hive Ranger policy with all permissions.

Resource Policies

Tag Policies

Reports

Audits

Security Zone

Settings

Create Policy

Service Manager > Hadoop SQL Policies > Create Policy

Last Response Time  
07/25/2024 05:11:52 PM

Policy Details

Policy Type

Access

Add Validity Period

Policy Name\*

Hive\_Ranger

Enabled

Normal

Policy Label

Hive

DB

Include

Hive Database

Hive Table

Hive Column

Include

Include

Include

Description

Hive Ranger Policy

Audit Logging\*

Yes

Allow Conditions:

hide

Select Roles

Select Groups

Select Users

Permissions

Delegate Admin

Select...


Select...

cdso\_security

select update Create Drop Alter Index Lock All Read Write ReplAdmin Service Admin Temporary UDF Admin Refresh RW Storage

Atlas Ranger policy

Additionally, you must set up Ranger policies for Atlas to work with asset search and tag flow management.

 **Note:** When a subscriber is provided with user access either as EnvironmentAdmin or EnvironmentUser, by default, the user is synchronized with the default Atlas policies in Ranger. This implies that the same user has access to all the datasets in the environment.

The following diagram provides a sample Atlas Ranger policy.

Resource Policies

Tag Policies

Reports

Audits

Security Zone

Settings

Create Policy

Service Manager > cm\_atlas Policies > Create Policy

Last Response Time  
07/25/2024 05:23:57 PM

Policy Details

Policy Type

Access

Add Validity Period

Policy Name\*

Policy Name

Enabled

Normal

Policy Label

Atlas

Ranger

Include

Type Category

Type Name

Include

Include

Description

Atlas Ranger Policy

Audit Logging\*

Yes

Allow Conditions:

hide

Select Roles

Select Groups

Select Users

Permissions

Delegate Admin

Select...

Select...

Read Type Create Type Update Type Delete Type

Reducing resource consumption with restricted users

When you plan to restrict data access, note the following:

- Audit summarization for the asset evolves from the Ranger Audit Profiler and Metrics service.
- Various Hive Column Statistical metrics for columns of the asset evolves from Atlas as part of the profile\_data of a column.

To ensure that the data related to audit summary and Hive Column Statistics are not visible to the subscribers, you must make sure to turn off the Ranger Audit Profiler/Activity Profiler and the Hive Column Profiler/Statistics Collector Profiler respectively.

5

### Related Information

[Providing role access](#)

[Ranger Policies Overview](#)

[Resource-based Services and Policies](#)

[Configuring resource-based policies](#)

## FAQs

This page provides answers to frequently asked questions about problems that you might encounter while you use Cloudera Data Catalog.

### Issue while accessing Cloudera Data Catalog

You must note about specific permissions that are needed to access Cloudera Data Catalog.

Users have to add themselves to at least one Data Lake as an Environmental Admin or Environmental User. This allows you to search for assets on the Data Lake, add tags, and business glossaries to these assets.

In order to use Profilers (Cloudera Data Hub clusters specific to Cloudera Data Catalog or Kubernetes based profiling jobs in Compute Cluster based environments) related features, you must have a PowerUser role. This runs Spark jobs and provides some valuable insights / suggestions.

## Associating tags to Hive assets

Associating tags to Hive assets fails on 7.2.16 cluster.



**Note:** This issue only applies for VM-based environments.

The issue can be resolved by adding the following Spark configuration:

`spark.sensitive.profiler.tableSamplePercentage=50` to profiler\_scheduler conf by adding the configuration under the key - `profiler_scheduler_spark_conf`.

The value passed in the configuration must be an integer value ranging from 1 to 100.

This value indicates the percentage of rows to be picked up as a sample for the profiler job.