

## Troubleshooting the Management Console

Date published: 2019-08-22

Date modified:



# 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

<b>Checking CDP logs.....</b>	<b>4</b>
Collecting Data Lake diagnostics via Cloudera Manager.....	4
Accessing Cloudera Manager logs.....	5
<b>Obtain CDP tenant ID.....</b>	<b>5</b>
<b>Credential creation errors on AWS.....</b>	<b>5</b>
<b>Credential creation errors on Azure.....</b>	<b>6</b>
<b>Specifying a group when user belongs to multiple groups.....</b>	<b>8</b>
<b>General networking and authentication errors.....</b>	<b>10</b>
<b>Problem deleting resources when using AWS Firewall Manager.....</b>	<b>10</b>

## Checking CDP logs

When troubleshooting Environment issues and Data Lake clusters, you can access the following logs:

### Collecting Data Lake diagnostics via Cloudera Manager

Cloudera Manager can collect system status, configuration details, logs, and other information from a Data Lake cluster into a zip file to send to Cloudera Support.

To help with solving problems when using Cloudera Manager on your data lake cluster, Cloudera Manager can collect diagnostic data. You can choose to trigger this collection to aid with resolving a problem, or you can configure Cloudera Manager to send a diagnostic bundle to Cloudera on a regular schedule. Cloudera Support analyses the diagnostic information to proactively identify problems.

To trigger creation of a diagnostic bundle for a Data Lake, navigate to the Cloudera Manager UI that manages the Data Lake cluster. In the lower section of the left navigation pane, choose **Support Send Diagnostic Data** and set the options in the dialog box that opens.

Specifically:

1. Log in to CDP web interface.
2. In the left navigation panel, click **Environments**, and then search for the appropriate environment and click the environment name.
3. In the environment detail page, open the **Data Lake** tab.
4. Open Cloudera Manager by clicking its link in the **Services** section.
5. In the bottom of the left navigation pane, click **Support Send Diagnostic Data**.
6. Choose whether to send the diagnostic bundle to Cloudera automatically or only to collect the bundle.

If the Cloudera Manager host does not have an internet connection, you may want to collect the bundle then move it to a host with access to the internet.

7. If appropriate, enter a Cloudera Support case number.
8. Select the Data Lake cluster if this instance of Cloudera Manager is managing more than one cluster.
9. If appropriate, limit the diagnostic collection to a single host, service, or role.

Open the **Restrict log and metrics collection** and choose the Data Lake host, service, or role for which you want to collect diagnostics.

10. Set the logic for what data to collect, by target size or by date range.

If you choose a target size, set the end time to a time to a few minutes after the event that you are trying to capture diagnostics for. The time range is based on the timezone of the host where Cloudera Manager Server is running.

11. Add a comment to describe the reason for collecting the diagnostic data.
12. Start the collection by clicking **Collect Diagnostic Data** or **Collect and Upload Diagnostic Data**.

The Cloudera Manager task dialog appears to track the jobs involved in collecting diagnostics.

13. When the collection tasks are complete, click **Download** to save the diagnostic bundle locally.

#### Related Information

[Diagnostic Data Collection in Cloudera Manager](#)

[Collecting Usage and Diagnostic Data](#)

## Accessing Cloudera Manager logs

CDP uses Cloudera Manager to orchestrate the installation of Data Lake components. Each instance in the cluster runs an Cloudera Manager agent which connects to the Cloudera Manager server. Cloudera Manager server is declared by the user during the cluster installation wizard.

Refer to the following documentation for information on how to access Cloudera Manager logs:

### Related Information

[Cloudera Manager logs](#)

## Obtain CDP tenant ID

Your CDP tenant ID can be found in the user and machine user profile pages, just below the resource CRN.

You may need to obtain your CDP tenant ID, for example when working with Cloudera support. Use the following steps to find your CDP tenant ID.

Steps

1. In CDP web interface, navigate to Management Console > User Management.
2. Find your user and click on it to navigate to details.
3. Your Tenant ID is listed among other user details.

## Credential creation errors on AWS

The following section lists common issues related to creating a credential on AWS and steps to resolve them.

### User: arn:aws:iam:::user/assume-only-user is not authorized to perform: sts:AssumeRole

Error: The following error occurs when creating a role-based AWS credential:

```
User: arn:aws:iam:::user/assume-only-user is not authorized to perform: sts:AssumeRole
```

Solution: The error occurs when CDP is not authorized to use the role that you are trying to register as part of cloud credential creation. The most common reason for this error is that when creating the cross-account IAM role you did not provide the AWS account ID. Refer to the documentation for creating a role-based credential on AWS for correct steps to create a cross-account IAM role.

### Internal error when creating an AWS credential from CDP CLI

Error: The following error occurs when creating a role-based AWS credential via CDP CLI:

```
An error occurred: An internal error has occurred. Retry your request, but if the problem persists, contact us with details by posting a message on the Cloudera Community forums. (Status Code: 500; Error Code: UNKNOWN; Service: environments; Operation: createAWSCredential; Request ID: dbf1fb6f-3161-46e1-80e9-a894461ceec3;)
```

Solution: A common reason for this error is that the external ID that you used for your cross-account IAM role is associated with another user. The external ID is tied to the CDP user who obtained it; Therefore, only the user who obtained the external ID is able to complete the credential creation flow in CDP with a given external ID. If a CDP user tries using an IAM role with an external ID obtained by another CDP user, CDP will return this error message.

To resolve the issue, log in to CDP CLI, obtain a new external ID, create a new cross-account IAM role, and then try creating the credential again.

### Related Information

[Create a cross-account IAM role](#)

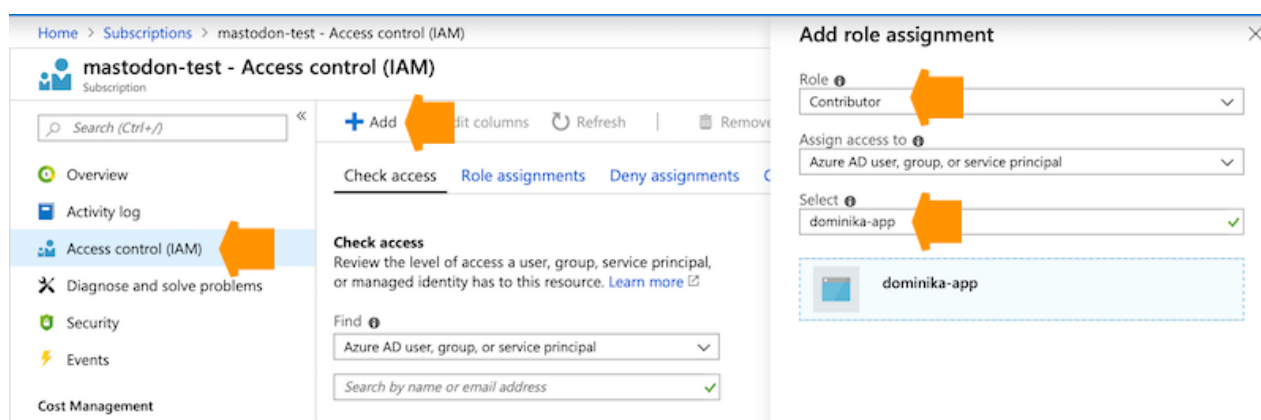
## Credential creation errors on Azure

The following section lists common issues related to creating a credential on Azure and steps to resolve them.

### You don't have enough permissions to assign roles

Error: You don't have enough permissions to assign roles, please contact with your administrator

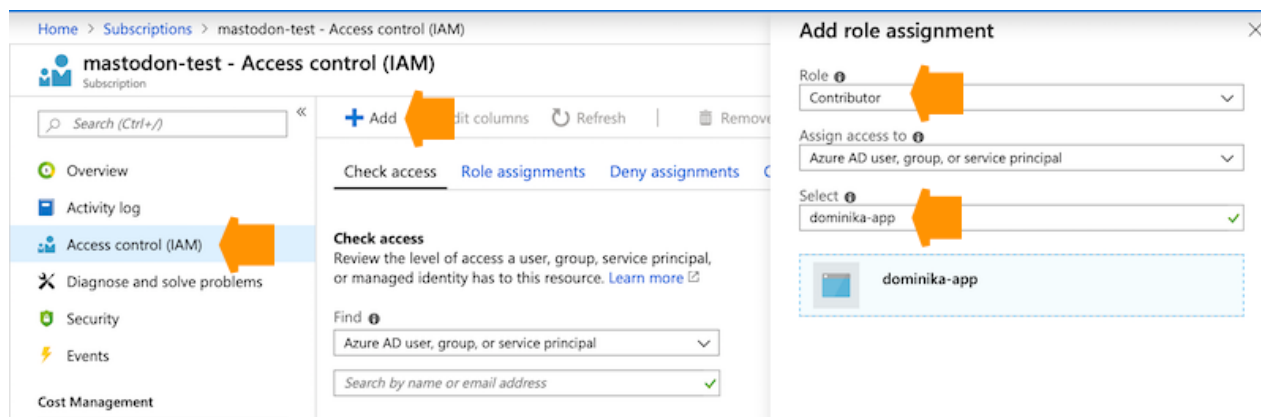
Solution: This error occurs in the Azure console because you are not authorized to perform role assignment. To solve the problem, ask your Azure administrator to perform the step of assigning the Contributor role to your application.



### Client does not have authorization

Error: Failed to verify credential: Status code 403, {"error":{"code":"AuthorizationFailed", "message":"The client 'X' with object id 'z' does not have authorization to perform action 'Microsoft.Storage/storageAccounts/read' over scope 'subscriptions/...'}}

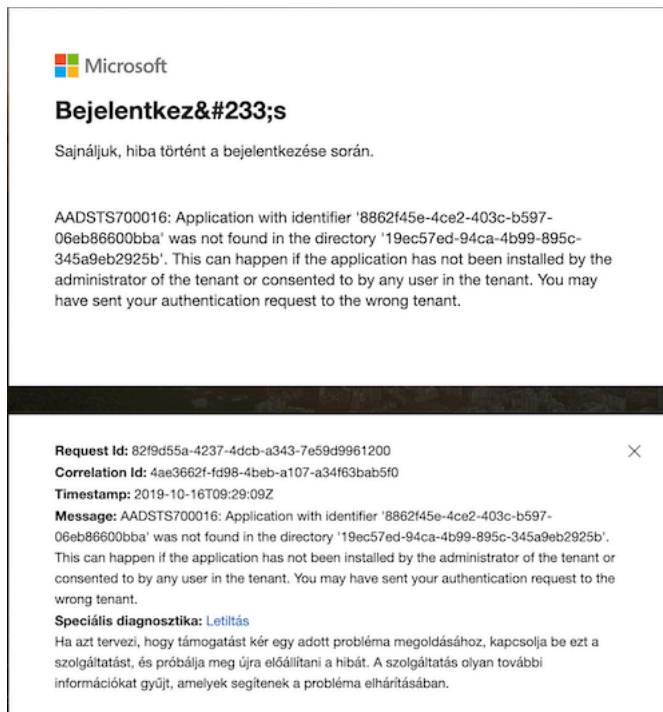
Solution: If you get this error during app-based credential creation, then the reason is that you did not assign the Contributor role to your app. To solve the issue, either assign the Contributor role to the app or - if you don't have Owner role - ask your Azure administrator to perform this step.



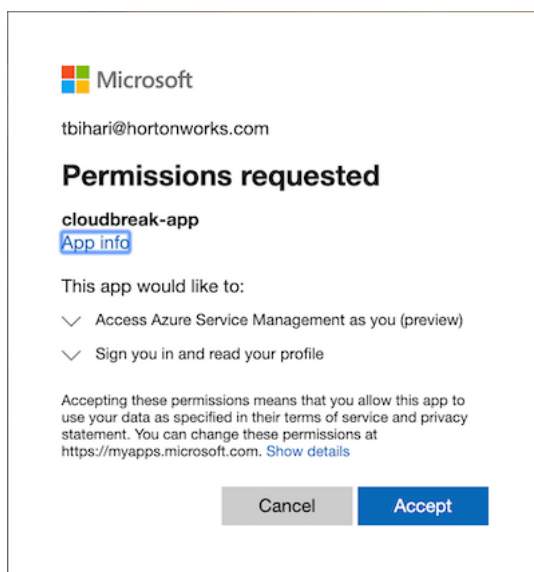
## Application with identifier was not found in the directory

Error: The following error appears during the interactive credential creation flow:

Application with identifier was not found in the directory. This can happen if the application has not been installed by the administrator of the tenant or consented to by any user in the tenant. You may have sent your authentication request to the wrong tenant.



Solution: To fix the issue, wait for one minute and then refresh the page. Afterwards, you should see the application consent page:



## Specifying a group when user belongs to multiple groups

If a user is mapped to multiple roles via group membership, the specific role to be used needs to be provided at runtime.

**Note:**

Spark does not work with Hive Metastore if there are multiple IDBroker group mappings. If you would like to use multiple group mappings, you should use the Ranger Authorization Service (RAZ).

If you do not specify the group, you will get an error similar to the following:

```
Ambiguous group role mappings for the authenticated user.
```

To provide the role, use the following commands:

```
fs.<s3a | azure | gs>.ext.cab.required.group
```

This command specifies the name of a particular group from which the credentials should be resolved. This can be used, when multiple user groups resolve to cloud roles, to specify for which group credentials are needed.

```
fs.<s3a | azure | gs>.ext.cab.required.role
```

This command specifies a particular cloud IAM role for which the credentials are needed. This can be used when the required cloud IAM role is known, and the IDBroker is configured in any way to resolve the user to that role.

The following examples illustrate how to use these commands in different contexts:

### HDFS CLI

The following command sets a Java system property (-D) called "fs.s3a.ext.cab.required.group" with the value of the group to use (GROUP).

```
hdfs dfs -Dfs.s3a.ext.cab.required.group=GROUP -ls s3a://BUCKET/PATH
```

The system property is used by the s3a binding to request that group specifically instead of falling back to looking up all the groups for the user.

### Spark

Here is how to set the config for a Spark job:

```
spark-shell --conf spark.yarn.access.hadoopFileSystems=s3a://BUCKET/ --conf
spark.hadoop.fs.s3a.ext.cab.required.group=GROUP

val textFile = spark.read.text("s3a://BUCKET/PATH")
textFile.count()
textFile.first()
```

### Livy

Here is how to set the config for a batch job:

```
# Ensure credentials
kinit

# Create test.py with following content
from pyspark.sql import SparkSession
```



```

spark = SparkSession.builder.appName("SimpleApp").getOrCreate()
textFile = spark.read.text("s3a://cldr-cdp-dl-1/user/systest/test.csv")
textFile.count()

# Upload file to cloud storage
hdfs dfs -copyFromLocal -f test.py s3a://cldr-cdp-dl-1/user/systest/test.py

# Submit Livy batch job
curl --negotiate -u: -i -XPOST --data '{"file": "s3a://cldr-cdp-dl-1/user/systest/test.py", "conf":{"spark.yarn.access.hadoopFileSystems":"s3a://cldr-cdp-dl-1/", "spark.hadoop.fs.s3a.ext.cab.required.group":"GROUP"}}' -H "Content-Type: application/json" http://krisden-1.gce.cloudera.com:8998/batches
# Check Livy batch job
curl --negotiate -u: -i http://krisden-1.gce.cloudera.com:8998/batches/0
curl --negotiate -u: -i http://krisden-1.gce.cloudera.com:8998/batches/0/log

```

Here is how to create a new session in Livy using the right config:

```

# Ensure credentials
kinit

# Create Livy session
curl --negotiate -u: -i -XPOST --data '{"kind": "pyspark", "conf":{"spark.yarn.access.hadoopFileSystems":"s3a://cldr-cdp-dl-1/", "spark.hadoop.fs.s3a.ext.cab.required.group":"GROUP"}}' -H "Content-Type: application/json" http://krisden-1.gce.cloudera.com:8998/sessions

# Check Livy session is idle
curl --negotiate -u: -i http://krisden-1.gce.cloudera.com:8998/sessions
curl --negotiate -u: -i http://krisden-1.gce.cloudera.com:8998/sessions/0
curl --negotiate -u: -i http://krisden-1.gce.cloudera.com:8998/sessions/0/state

# Use Livy session
curl --negotiate -u: -i -XPOST --data '{"kind": "pyspark", "code": "spark.read.text(\"s3a://cldr-cdp-dl-1/user/systest/test.csv\").count()"}' -H "Content-Type: application/json" http://krisden-1.gce.cloudera.com:8998/sessions/0/statements

# Check the session statement output
curl --negotiate -u: -i http://krisden-1.gce.cloudera.com:8998/sessions/0/statements/0

# Delete the session
curl --negotiate -u: -i -XDELETE http://krisden-1.gce.cloudera.com:8998/sessions/0
curl --negotiate -u: -i http://krisden-1.gce.cloudera.com:8998/sessions

```

## Zeppelin

Make sure to set the following configs in the Livy interpreter config before running the notebook:

- Set livy.spark.yarn.access.hadoopFileSystems to s3a://cldr-cdp-dl-1/
- Set livy.spark.hadoop.fs.s3a.ext.cab.required.group to GROUP

Next, create a Zeppelin notebook with the following content:

```

%livy.pyspark
textFile = spark.read.text("s3a://cldr-cdp-dl-1/user/systest/test.csv")
textFile.count()

```

**Note:**

Livy interpreter uses Livy sessions and not batches.

## General networking and authentication errors

Are you finding that your CDP interactions are hitting errors when they shouldn't? Check to see if your identity management system is working. FreeIPA errors can be hidden as failures from Knox, Kerberos, or other systems that act as clients to FreeIPA.

You can eliminate FreeIPA issues by checking the following status:

- Management Console: Select the Environment Summary FreeIPA status .

ID	Instance Status	Status Reason	FQDN ↑	Private IP	Public IP
<a href="#">0929c15e13ac6f0fa</a>	Unreachable	missing data	ipaserver0.jamison.xcu2-8y8x.wl.cloudera.site	10.117.232.31	N/A
<a href="#">0d470b66dfe3abadf</a>	Deleted on provider	missing data	ipaserver1.jamison.xcu2-8y8x.wl.cloudera.site	10.117.234.77	N/A
<a href="#">000d9d82278a0f326</a>	Running	missing data	ipaserver2.jamison.xcu2-8y8x.wl.cloudera.site	10.117.234.253	N/A

- Cloud provider instance status: Follow the link for one of the FreeIPA hosts in the status to open the cloud provider's status page for that host. If one of the FreeIPA hosts is stopped, restart it. It is possible to still see a "Running" status for FreeIPA even if one of the hosts is stopped.

### Related Information

[Managing FreeIPA](#)

## Problem deleting resources when using AWS Firewall Manager

If you have AWS Firewall Manager enabled, ensure that your AWS Firewall Manager policies or rules don't affect any AWS resources created by CDP. If they do, you may experience problems deleting these resources due to dependency violation.

For more information, refer to [Security group policies](#) > Avoid conflicts if you also use outside sources to manage security groups in AWS documentation.