

Cloudera Data Engineering 1.5.4

Creating and managing CDE Sessions

Date published: 2020-07-30

Date modified: 2024-05-30

The Cloudera logo is displayed in a bold, orange, sans-serif font. The word "CLOUDERA" is written in all caps, with a stylized 'E' that has a horizontal bar extending to the right.

<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

Creating Sessions in Cloudera Data Engineering.....	4
Interacting with a Session in Cloudera Data Engineering.....	5
Connecting Sessions with the CDE CLI.....	6
Viewing logs for Cloudera Data Engineering Sessions.....	6
Viewing Spark UI for Cloudera Data Engineering Sessions.....	6

Creating Sessions in Cloudera Data Engineering

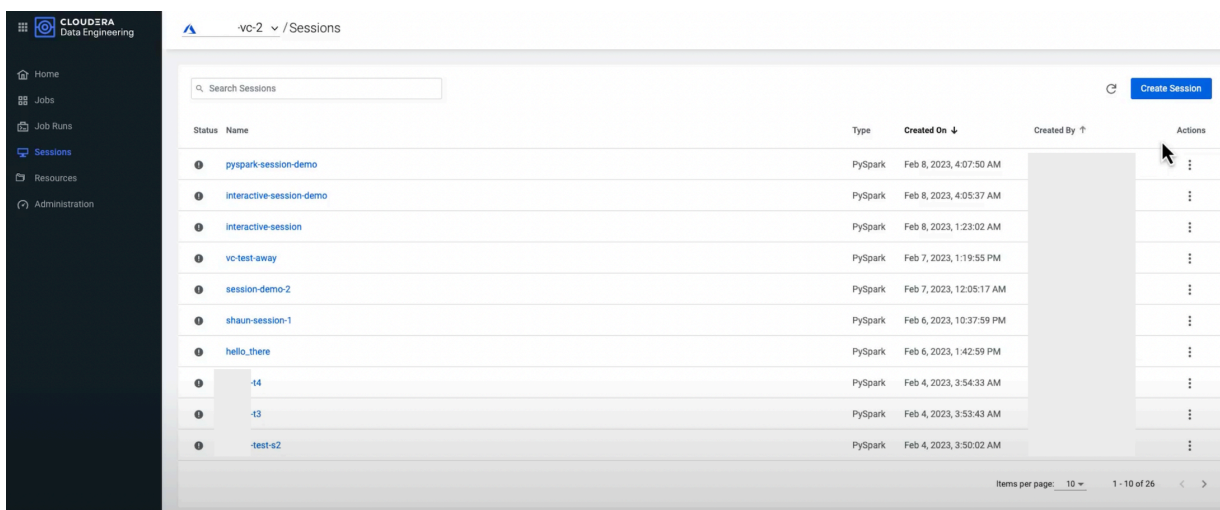
A Cloudera Data Engineering (CDE) Session is an interactive short-lived development environment for running Spark commands to help you iterate upon and build your Spark workloads.

About this task

The commands that are run in a CDE Session are called Statements. You can submit the Statements through the connect CLI command or the Interact tab in the CDE UI for a Session. Python and Scala are the supported Session types. Learn how to use Cloudera Data Engineering (CDE) Sessions using the user interface and CLI.

Procedure

1. In the Cloudera Data Platform (CDP) console, click the Data Engineering tile. The Home page displays.
2. Click Sessions in the left navigation menu and then click Create Session.



3. Enter a Name for the Session.
4. Select a Type, for example, PySpark or Scala.
5. Select a Timeout value.
The Session will stop after the indicated time has passed.
6. Optionally, enter a Description for the Session.
7. Optionally, enter the Configurations.
8. Set the Compute options.
 - Optional: GPU Acceleration (Technical Preview): You can accelerate your session using GPUs. Click Enable GPU Accelerations checkbox to enable the GPU acceleration and configure selectors and tolerations if you want to run the job on specific GPU nodes. When you run this session, this particular session will request GPU resources.

- Under **Files and Resources**, you can upload Jar, Python, Egg, Zip, and other files. You can also add a resource, repositories, or a Python environment to be used in this session.

Files that are uploaded to a session are stored in the app/mount directory.

- Click Create.

The Connect tab displays a list of connectivity options available to interact with the Session. The Interact tab allows you to interact with the Session, and becomes available once the Session is running.

- To delete a Session, open the Session and click Delete.



Note: If you delete a Session, doing so will result in the termination of an active session and the loss of any attached logs and details.

Interacting with a Session in Cloudera Data Engineering

Once a Session is created in Cloudera Data Engineering (CDE), you can begin your interaction.

About this task

You can also interact with files that were uploaded to a Session. Those files are stored in the app/mount directory.

Procedure

- In the Cloudera Data Platform (CDP) console, click the Data Engineering tile. The Home page displays.

2. Click Sessions in the left navigation menu.
3. Click the Interact tab.
4. In the Input box, enter your commands.
5. Click Run. A history of executed commands displays in the sequence of their execution.

Connecting Sessions with the CDE CLI

Once a Session is created in Cloudera Data Engineering (CDE), you can download the CDE CLI from the user interface then use the terminal on your local machine to set up the interactive session using the CDE CLI.

Procedure

1. In the Cloudera Data Platform (CDP) console, click the Data Engineering tile. The Home page displays.
2. Click Sessions in the left navigation menu.
3. Click the Connect tab.
4. Download the CDE CLI.
5. Use the CDE CLI to run and connect interactive commands:
 - a) Copy the command that displays.
 - b) Paste the command in the CDE CLI and continue to interact with your session.

Viewing logs for Cloudera Data Engineering Sessions

Once a Session is created in Cloudera Data Engineering (CDE), you can view logs of the Spark driver and executors from the Session.

Procedure

1. In the Cloudera Data Platform (CDP) console, click the Data Engineering tile. The Home page displays.
2. Click Sessions in the left navigation menu.
3. Click the Session name that you want to view the logs for.
4. Click the Logs tab.
5. In the Select Log Type drop-down, select the Driver or Executor that you want to view the logs for.
6. Click Download All Logs to save your log files.

Viewing Spark UI for Cloudera Data Engineering Sessions

You can view the Spark UI for a Cloudera Data Engineering (CDE) active Session for troubleshooting purposes.

Procedure

1. In the Cloudera Data Platform (CDP) console, click the Data Engineering tile. The Home page displays.
2. Click Sessions in the left navigation menu.
3. Select the Virtual Cluster from the drop-down menu at the top.
4. Select the Active/Running sessions that you want to troubleshoot.
5. Click the Spark UI tab.