

Managing Flow Definitions

Date published: 2021-04-06

Date modified: 2024-06-03



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

Managing flow definitions using the Catalog.....	4
Accessing flow definitions.....	4
Management of flow catalog.....	5
Management of flow definition details.....	6
Adding custom tags to flow definitions.....	8
Downloading the latest version of a flow definition.....	8
Deleting a flow definition.....	9

Managing flow definitions using the Catalog

In Cloudera DataFlow, your flow definitions are available in the Catalog. You can browse, discover, and manage flow definition versions in this view. You can also deploy and reuse your flow definitions across multiple environments.

Before you begin

- Make sure that you have DFCatalogAdmin permission. For information on account and resource roles, see *Cloudera DataFlow Authorization*.
- Import your Apache NiFi flow into Cloudera DataFlow, or publish your draft from Flow Designer to make it available as a flow definition in the Catalog.

Related Information

[Cloudera DataFlow Authorization](#)

[Importing a flow definition from NiFi](#)

[Publishing a draft as a flow definition from Flow Designer](#)

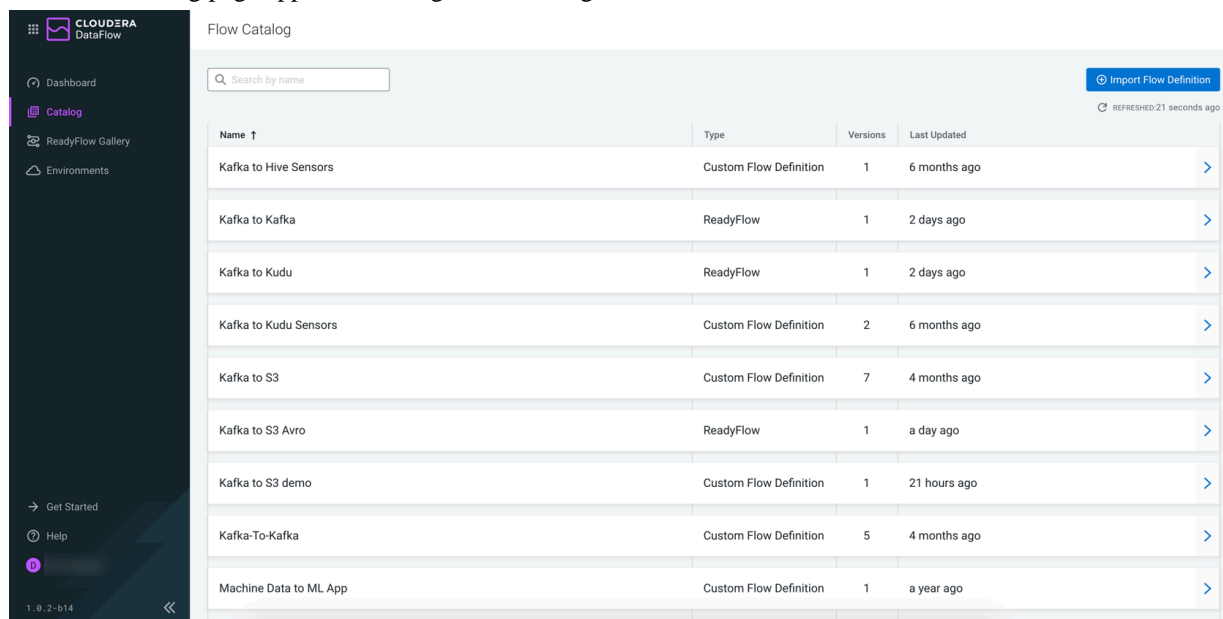
Accessing flow definitions

Learn how to access the Catalog page and flow definition details pane for managing flow definition details.

Procedure

1. Go to the Cloudera DataFlowUI.
2. Click Catalog from the left navigation pane.

The Flow Catalog page appears showing the following information.



Name ↑	Type	Versions	Last Updated
Kafka to Hive Sensors	Custom Flow Definition	1	6 months ago
Kafka to Kafka	ReadyFlow	1	2 days ago
Kafka to Kudu	ReadyFlow	1	2 days ago
Kafka to Kudu Sensors	Custom Flow Definition	2	6 months ago
Kafka to S3	Custom Flow Definition	7	4 months ago
Kafka to S3 Avro	ReadyFlow	1	a day ago
Kafka to S3 demo	Custom Flow Definition	1	21 hours ago
Kafka-To-Kafka	Custom Flow Definition	5	4 months ago
Machine Data to ML App	Custom Flow Definition	1	a year ago

For more information about what to track in the Flow Catalog page, see *Management of flow catalog*.

- Click a flow definition row to expand and view additional information about it.

The Flow Definition Details pane appears showing the following information.

The screenshot shows the Cloudera DataFlow interface. On the left is a sidebar with navigation links: Dashboard, Catalog, ReadyFlow Gallery, Flow Design, Projects, Functions, and Environments. The main area is divided into two panes. The left pane, titled 'Flow Catalog', lists several flow definitions: gcs-to-gcs-processing-via-designer-vett, Generate Flow File Log, Google Drive to S3/ADLS, Hello World, Hello World CSV (selected), Hello World Publish to Kafka, Hello World2, Hive Ingest, HTTP Gateway to Kafka, HTTP Gateway to Kafka and SnowFlake, HTTP Gateway to Kafka and SnowFlake 2, and HubSpot to S3/ADLS. The right pane, titled 'Hello World CSV', shows details for this flow definition. It includes a 'Refreshed: 7 seconds ago' indicator, an 'Actions' dropdown, and a checkbox for 'Only show deployed versions'. Below this is a table with columns: Version, Deployments, Drafts, and Tags. The table shows two versions: Version 2 (0 deployments, 0 drafts, Tech Preview tag) and Version 1 (1 deployment, 0 drafts, Obsolete tag). Below the table are buttons for 'Deploy', 'Download', and 'Create New Draft'. The 'DEPLOYMENTS (1)' section shows a deployment named 'cdf-priv-azure' with a sub-item 'Deployment-1'. The 'CRN #' section shows a long alphanumeric string. The 'CREATED' section shows the date '2023-02-28 16:24 GMT+1' by Miklos Kertesz. The 'Custom Tag' section shows a text input field and a 'Color' dropdown menu.

For more information about what to track in the Flow Definition Details pane, see *Management of flow definition details*.

Related Information

[Management of flow catalog](#)

[Management of flow definition details](#)

Management of flow catalog

Learn the information that you can track in the Catalog page. Also, learn how to sort, search, and refresh your flow definitions.


You can track the following details of flow definitions in the Catalog page:

- Name – Name of the flow definition
- Type – Type of flow definition, Custom or ReadyFlow
- Versions – Number of available versions
- Last Updated – When the flow definition was last changed

You can sort flow definitions by Name, Type and Last Updated. Click the column name to sort by ascending or descending order.

You can also use the Search field at the top of the **Flow Catalog** page to directly search for the flow definition by its name.

You can use the Items per page option, at the bottom of the screen, to display the desired number of flow definitions in your page view.

You can click the  Refresh button in the top right corner to refresh the data. By default, every 30 seconds all data for your flow definitions are updated.

Management of flow definition details

Find out about information that you can track in the Flow Definition Details pane. Also, learn how to import new version, download latest version, and delete your flow definitions.

You can track the following details of flow definitions in the Flow Definition Details pane:

- Flow description – This field gives you additional information about the flow definition.



Note: You add this description when you import the flow from NiFi, or when you publish it to the Catalog from Flow Designer. You cannot edit it afterwards.

- CRN # – This field allows you to copy the CRN of the flow definition
- Only show deployed versions – Select this option if you want to view only the deployed versions of a flow definition.

- List of flow definition versions - This table provides you information about specific versions of the flow definition. It also allows you to perform various actions, such as deploy, download, or create a draft using a specific version of the flow definition.

☐ Only show deployed versions

Version	Deployments	Drafts	Tags	
2	0	0	Tech Preview	▼
1	1	0	Obsolete	▲

Version

This column shows the version number of the flow definition. You can store multiple versions of a flow definition in the Flow Catalog.

Deployments

This column shows the number of deployments that belong to a specific version of a flow definition.

Drafts

This column shows the number of drafts that are based on a specific version of a flow definition.

Tags

This column shows any tags added to a specific version of a flow definition.

Click a flow definition version row to perform actions using that version.

Deploy

Select to deploy the specific version of the flow definition.

Download

Select to download the JSON file of the specific version of a flow definition. This option is not available for the ReadyFlow type of flow definitions.

Create New Draft

Select to create a new draft in Flow Designer based on that specific version of the flow definition.

CRN

This field enables you to copy the CRN number of the specific flow definition version.

Created

This field provides information about when the flow definition version was created and by which user. This section also includes the version comments entered by the user when the flow definition version was added.

Custom Tag

This field allows you to add free text tags to a specific version of a flow definition.

Deployments

This field shows you the number of deployments that belong to a flow definition. If you have the DFFlowAdmin role for the environments where the flow definition has been deployed, you can also see the list of the deployments.

- **Actions** – Use this drop-down menu to import a new version or delete a flow definition.

Import New Version

Select to upload new versions of an existing flow definition to the Flow Catalog.

Download Latest Version

Select to download the JSON file of the latest version of a flow definition from the flow catalog.

Delete Flow Definition

Select to remove a flow definition from the flow catalog along with all its versions.



Adding custom tags to flow definitions

You can add informative tags to individual versions of a flow definition in the Catalog.

About this task

Custom tags allow you to dynamically introduce additional information to flow versions on top of the static, non-modifiable version information. Once added, custom tags cannot be modified. You can delete an existing tag and replace it with a new one. If you change your mind, you can reintroduce a recently deleted tag using the Add Recently Used option.

Procedure

1. In the Cloudera DataFlow UI, from the left navigation pane, click Catalog.
Flow definitions available to you are displayed, one definition per row.
2. Identify the flow definition where you want to add tags to a flow version, and click the row to display the flow definition details and versions.
3. Expand the flow version where you want to add tags by clicking  Expand
4. Type the tag text under Custom Tag.
5. You can optionally change the default tag color, using the Color menu.
6. Click  Add.

Downloading the latest version of a flow definition

You can directly download the latest version of a flow definition from the Cloudera DataFlow UI.

Before you begin

You have imported at least one flow definition.

Procedure



1. In the Cloudera DataFlow UI, from the left navigation pane, click Catalog.
Flow definitions available for you are displayed, one definition per row.
2. Identify the flow definition you wish to download, and click the row to display the flow definition details and versions.
3. From the Actions drop-down, select Download Latest Version.
4. Click Download.

Results

The JSON file of the flow definition is downloaded. After you download the JSON file for a flow definition, you need to import it to add it to the Flow Catalog.

Example

» REFRESHED: 15 seconds ago


 Covid Data Stream
Updated 6 months ago by 

Actions ▾

- Import New Version
- Download Latest Version**
- Delete Flow Definition

FLOW DESCRIPTION

This flow reads covid data from several sources and writes it to CDP

CRN #
 crn:cdp:df:us-west-1:9d74eee4-1cad-45d7-b645-7ccf9edbb73d:flow:Covid-Data-Stream

☐ Only show deployed versions

Version	Deployments	
13	0	▾
12	0	▾

Deleting a flow definition

You can delete a flow definition along with all its versions from the Cloudera DataFlow UI.

Before you begin

You have imported at least one flow definition.

Procedure


1. In the Cloudera DataFlow UI, from the left navigation pane, click Catalog.
Flow definitions available for you are displayed, one definition per row.
2. Identify the flow definition you wish to delete, and click the row to display the flow definition details and versions.

- 3. From the Actions drop-down, select Delete Flow Definition.

Delete Flow Definition

Warning

Deleting a flow definition will remove all versions of it from the flow catalog.

FLOW DEFINITION NAME	VERSIONS	DEPLOYMENTS
Kafka to Hive Sensors 	1	0

Please enter the flow definition name to confirm:

Cancel

Delete

- 4. Enter the flow definition name, and click Delete.

10