

Cloudera DataFlow for Data Hub 7.2.18

Flow Management Overview

Date published: 2019-12-16

Date modified: 2024-04-03

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

What is Apache NiFi?.....	4
What is Apache NiFi Registry?.....	4
Cloudera Manager integration.....	4

What is Apache NiFi?

Cloudera Flow Management (CFM) is a no-code data ingestion and management solution powered by Apache NiFi.

CFM includes two primary components, NiFi and NiFi Registry. NiFi serves as the core data ingestion engine, with a wide range of processors, facilitating seamless connectivity, transformation, and content routing. It enables easy data ingestion and movement between CDP services, as well as 3rd party cloud services. NiFi has a no-code graphical user interface and offers a user-friendly environment to design and manage data flows.

For more information about the fundamental concepts behind NiFi, see the *Apache NiFi Overview* available in the Apache NiFi documentation.



Important:

The Apache NiFi documentation is available both online and within the NiFi user interface (UI). The online documentation consistently reflects the latest NiFi version. If you are using an earlier NiFi version, content differences may exist between the latest documentation and the version you would need. You can access the documentation specific to your NiFi version directly from the NiFi UI:

1. Launch the NiFi UI.
2. From the Global Menu, select `Help Documentation`.

Related Information

[Apache NiFi Overview](#)

What is Apache NiFi Registry?

Cloudera Flow Management (CFM) is a no-code data ingestion and management solution powered by Apache NiFi.

CFM includes two primary components, NiFi and NiFi Registry. NiFi Registry, a companion to Apache NiFi, enables DevOps style development and deployment of flow files. It also supports flow versioning, promoting flows from one environment into another, and flow deployment. It is automatically set up when you create the Flow Management cluster and it serves as a centralized hub for managing your versioned data flows.

For more information about the fundamental concepts behind NiFi Registry, see the *Getting Started with Apache NiFi Registry* guide in the Apache NiFi Registry documentation.



Important:

The Apache NiFi Registry documentation is available both online and within the NiFi Registry user interface (UI). The online documentation consistently reflects the latest NiFi Registry version. If you are using an earlier NiFi Registry version, content differences may exist between the latest documentation and the version you would need. You can access the documentation specific to your NiFi Registry version directly from the NiFi Registry UI:

1. Launch the NiFi Registry UI.
2. From the Global Menu, select `Help Documentation`.

Related Information

[Getting Started with Apache NiFi Registry](#)

Cloudera Manager integration

The NiFi service of Cloudera Flow Management (CFM) is integrated with Cloudera Manager on CDP Private Cloud Base.

Cloudera Manager monitoring

Cloudera Manager is retrieving monitoring data from NiFi by calling the below endpoint:

```
/nifi-api/flow/metrics/json?rootFieldName=beans&sampleName=.*_jvm_.*&sampleLabelValue=RootProcessGroup
```

Metrics and graphs for the NiFi service are available starting with version CFM 2.1.4 and it requires Cloudera Manager 7.4.5 or higher.

This endpoint is called by the Cloudera Manager agent running on each NiFi node and the identity of the node is used to make the call. If Ranger is used for authorization, the calls are authorized as the nodes belong to the nifi group.

Cloudera Manager alerting

Multiple health checks are configured out of the box for the NiFi service in Cloudera Manager:

- **NiFi node connection**

It checks if the NiFi node is disconnected or not from the NiFi cluster. This can help to detect cluster instability in case the NiFi node is up and running but the node is not part of the NiFi cluster. The check executes the NiFi CLI `get -nodes` command on each NiFi node, which calls the corresponding NiFi API. If Ranger is used for authorization, the health check is authorized as the nodes belong to the nifi group.

- **NiFi Content repository usage**

It is used to check how much disk space is used on the Content repository.

- **NiFi Provenance repository usage**

It is used to check how much disk space is used on the Provenance repository.

- **NiFi FlowFile repository usage**

It is used to check how much disk space is used on the FlowFile repository.