

Cloudera Flow Management Operator 2.8.0

CFM Operator Overview

Date published: 2024-06-11

Date modified: 2024-06-11

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

CFM Operator overview.....	4
-----------------------------------	----------

CFM Operator overview

Cloudera Flow Management (CFM) Operator allows you to deploy and manage NiFi clusters and NiFi Registry instances on Kubernetes application platforms. CFM Operator simplifies data collection, transformation, and delivery across your enterprise. Leveraging containerized infrastructure, the operator streamlines the orchestration of complex data flows.

Apache NiFi is the core data ingestion engine of the operator. NiFi is a robust tool for building data movement pipelines using a visual flow designer. With hundreds of built-in processors, it facilitates easy connections to various applications and allows for the transformation of data structures or formats as needed. Supporting both structured and unstructured data for streaming and batch integrations, NiFi is a fundamental component of modern data pipelines and is essential in multimodal Generative AI use cases.

Key capabilities

Deploying and managing NiFi

- Deploy NiFi easily and fast.
- Manage multiple NiFi deployments.
- Run different NiFi versions within each cluster (NiFi 1.25, 2.0 [Technical Preview]).
- Run NiFi in a single-node mode or as a distributed cluster.

Deploying and managing NiFi Registry

- Deploy one or more NiFi Registries.
- Connect a NiFi Registry instance to one or more NiFi clusters and use it for central flow versioning.

cfmctl command line tool for installation

- Simplify the installation and deployment of CFM Operator and NiFi and NiFi Registry CRDs.
- Quickly assess environment readiness.

NiFi clusters without ZooKeeper

- Enable NiFi deployment without ZooKeeper by handling leader election and state management directly within Kubernetes.
- Store NiFi cluster state within Kubernetes configuration maps.

Scaling NiFi clusters up and down

- Manually scale up/down your NiFi deployments by adjusting the number of NiFi nodes in the configuration file.
- Utilize Kubernetes Horizontal Pod Autoscaler (HPA) to monitor Central Processing Unit (CPU) utilization and trigger autoscaling based on resource consumption.

Data durability

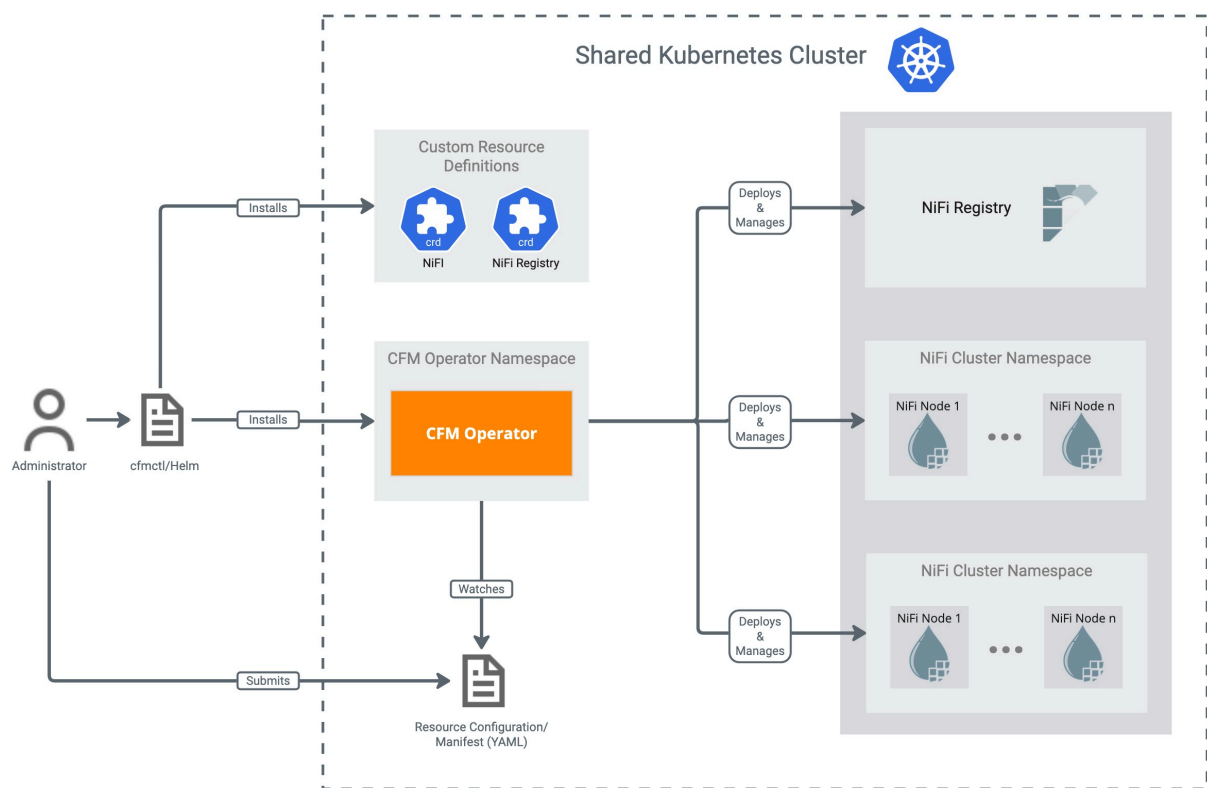
- In-flight data remains uninterrupted even if a NiFi node goes down. Kubernetes will provision a new node and transfer the volumes from the failed node to the new one, ensuring continuous data processing.

Full lifecycle management

- Change your configuration by deploying a modified configuration file and let the operator reconcile.
- Perform NiFi version upgrades effortlessly by specifying a new NiFi image. Support for rolling upgrades is coming soon for NiFi 2 clusters.

Deployment architecture

CFM Operator is responsible for managing NiFi and NiFi Registry deployments. It is deployed within a designated operator namespace, while the actual NiFi and NiFi Registry instances are managed within one or more separate namespaces. The following diagram shows a typical NiFi deployment with CFM Operator.



Licensing

CFM Operator requires a valid and active license to function. Licenses are made available to you together with your Cloudera credentials as part of your license and subscription agreement with Cloudera.

Licenses are registered during the installation of the CFM Operator and stored in a Kubernetes secret. They can be updated at any time.