

Scaling up or down a NiFi Cluster

Date published: 2019-12-16

Date modified: 2023-06-27

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.

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

Scaling up or down a NiFi cluster.....	4
Scaling up a NiFi cluster.....	4
Scaling down a NiFi cluster.....	5

Scaling up or down a NiFi cluster

You can scale up or down a NiFi cluster based on your requirement by adding or removing nodes respectively.

Scaling up a NiFi cluster

You have a NiFi cluster and you want to increase the throughput by adding new nodes to your NiFi cluster.

Before you begin

You have a NiFi flow which is running.

Procedure

1. Log into CDP Public Cloud UI.
2. Go to Management Console Environments .
3. Select the datahub cluster where your NiFi flow is running.

The screenshot shows the Cloudera Management Console interface. The left sidebar contains navigation options like Dashboard, Environments, Data Lakes, User Management, Data Hub Clusters, Data Warehouses, ML Workspaces, Classic Clusters, Cost Management, and Global Settings. The main content area displays details for a cluster named 'test110'. Key information includes:

- STATUS:** Running
- NODES:** 4
- CREATED AT:** 07/29/21, 08:16 PM GMT+2
- CLUSTER TEMPLATE:** 7.2.10 - Flow Management Light Duty with Apache NiFi, Apache NiFi Registry
- Environment Details:** NAME: vbrodetskiy-aws, DATA LAKE: vbrodetskiy-dl, CREDENTIAL: eng-dim-daily, REGION: us-west-2, AVAILABILITY ZONE: us-west-2c
- Services:** CM-UI, NiFi, NiFi Registry, Token Integration
- Cloudera Manager Info:** CM VERSION: 7.4.2, RUNTIME VERSION: 7.2.10-1.cdh7.2.10.p0.14376970
- Event History:** Shows events such as 'Removed instance(s)' and 'Removing instance(s)'. A 'DOWNLOAD' button is visible.

If you go to the NiFi UI by clicking NiFi, you can check the current number of NiFi nodes in your cluster, as shown in the following image:

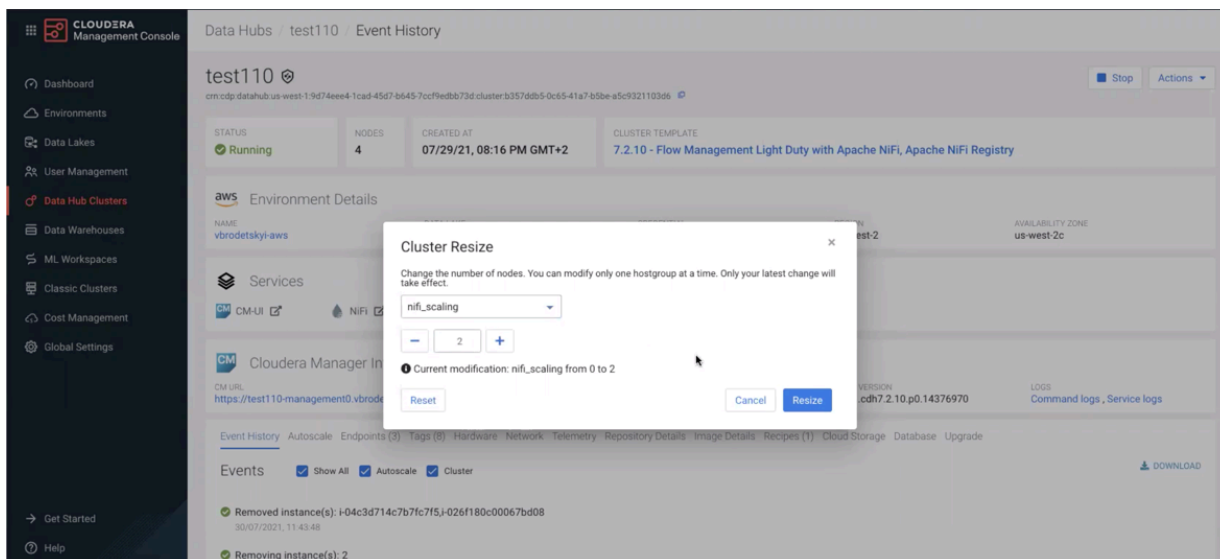
The screenshot shows the Cloudera Flow Management interface. The top navigation bar includes icons for various services. A red box highlights the 'NiFi' icon, which displays '3 / 3', indicating the current number of nodes in the cluster. Below the navigation bar, there is a 'Navigate' panel with search and view controls. The main workspace is a grid area with a 'GenerateFlow' button in the bottom right corner.

4. Click Actions, and select Resize.
The Cluster Resize window appears.

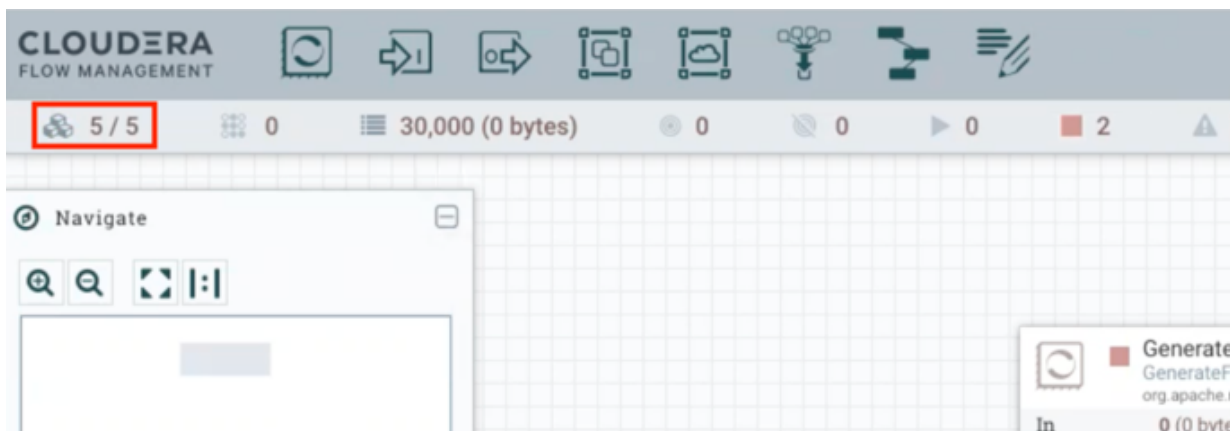
- Select your hostgroup from the dropdown, increase the number of nodes by clicking the + icon, and then click Resize.



Note: You need to select the hostgroup carefully. For light duty templates, the hostgroup is nifi_scaling. For heavy duty templates, the hostgroup is nifi.



It takes some time for the change to take place. In the NiFi UI you can see when the upscale event is completed and the new nodes are added to the NiFi cluster.



Results

You have now successfully scaled up your NiFi cluster.

Scaling down a NiFi cluster

You can also remove nodes from your NiFi cluster for cost optimization. For example, scale down the NiFi cluster when there is less data to be processed in order to save infrastructure costs.

Before you begin

You have a NiFi flow which is running.

Procedure

- Log into CDP Public Cloud UI.
- Go to Management Console Environments .

3. Select the datahub cluster where your NiFi flow is running.

The screenshot shows the Cloudera Management Console interface. The left sidebar contains navigation options like Dashboard, Environments, Data Lakes, User Management, Data Hub Clusters, Data Warehouses, ML Workspaces, Classic Clusters, Cost Management, and Global Settings. The main content area displays details for a cluster named 'test110'. Key information includes:

- STATUS:** Running
- NODES:** 6
- CREATED AT:** 07/29/21, 08:16 PM GMT+2
- CLUSTER TEMPLATE:** 7.2.10 - Flow Management Light Duty with Apache NiFi, Apache NiFi Registry

 Below this, there are sections for 'aws Environment Details' (listing name, data lake, credential, region, and availability zone), 'Services' (listing CM-UI, NiFi, NiFi Registry, and Token Integration), and 'Cloudera Manager Info' (listing CM URL, CM version, runtime version, and logs). At the bottom, a 'Management' table shows a list of nodes with columns for ID, FQDN, Status, Private IP, and Public IP. One node is visible with ID '+0edd329722b8a4390' and status 'Running'.

Before scaling down you can go to the NiFi UI by clicking NiFi to check the current number of NiFi nodes you have in your NiFi cluster, as shown in the following image:

The screenshot shows the Cloudera Flow Management NiFi UI. At the top, there is a toolbar with various icons. Below the toolbar, a status bar displays '5 / 5' nodes, indicating the current number of nodes in the cluster. The main area shows a flow diagram with a 'GenerateFlow' node. A tooltip for this node is visible, showing statistics:

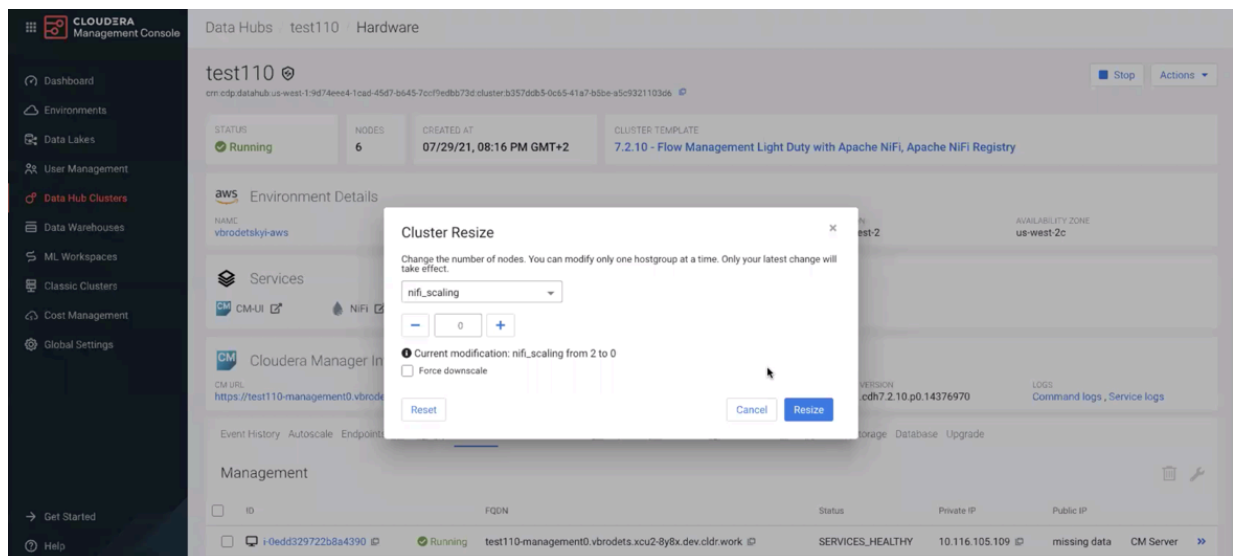
- In:** 0 (0 bytes)
- Read/Write:** 0 bytes / 0 I
- Out:** 20,000 (0 b)
- Tasks/Time:** 20,000 / 00

4. Click Actions, and select Resize.
The Cluster Resize window appears.

- Select your hostgroup from the dropdown, decrease the number of nodes by clicking the - icon, and then click Resize.



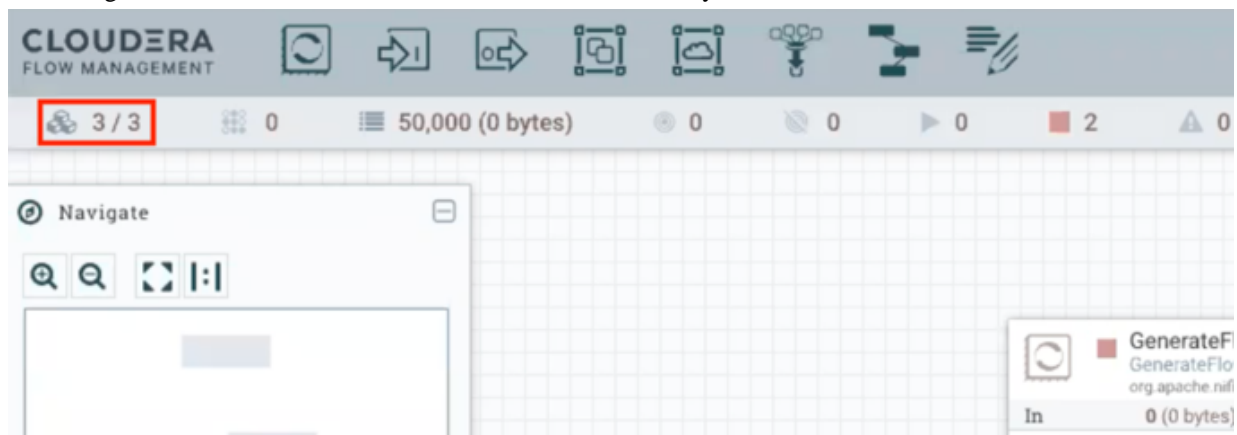
Note: You need to select the hostgroup carefully. For light duty templates, the hostgroup is nifi_scaling. For heavy duty templates, the hostgroup is nifi.



It takes some time for the change to take place. In the background, the following processes occur:

- The nodes to be removed are disconnected from the NiFi cluster.
- Processing of data on these nodes is stopped.
- The data located on the nodes to be removed are moved to the remaining nodes of the cluster. This ensures no data loss during a scale down event.
- The NiFi service is stopped on these nodes
- The underlying virtual machines are stopped and deleted.

You can go to the NiFi UI to check that the nodes are successfully removed from the NiFi cluster:



The scale down event is complete only after the nodes are no longer listed in the Hardware tab of the Data Hub Clusters page, as shown in the following image:

The screenshot shows the Cloudera Management Console interface. The left sidebar contains navigation options like Dashboard, Environments, Data Lakes, User Management, Data Hub Clusters, Data Warehouses, ML Workspaces, Classic Clusters, Cost Management, and Global Settings. The main content area is titled 'Data Hubs / test110 / Hardware' and shows 'Cloudera Manager Info' with details like CM URL, CM VERSION (7.4.2), and SOFTWARE VERSION (7.2.10-1.cdh7.2.10.p0.14376970). Below this are tabs for Event History, Autoscale, Endpoints (3), Tags (6), Hardware, Network, Telemetry, Repository Details, Image Details, Recipes (1), Cloud Storage, Database, and Upgrade. The 'Management' section shows a table with columns for ID, FQDN, Status, Private IP, and Public IP. The 'NiFi' section shows three running nodes with status 'SERVICES_HEALTHY'. The 'NiFi_scaling' section shows two decommissioned nodes with status 'DECOMMISSIONED'.

ID	FQDN	Status	Private IP	Public IP
+0ed6929722b6a4390	test110-management0.vbrodets.xcu2-8yflx.dev.cldr.work	Running	10.116.105.109	missing data
+0107e5867a0633cde	test110-nifi1.vbrodets.xcu2-8yflx.dev.cldr.work	Running	10.116.107.188	missing data
+07a7a3a02a5f0ae2c	test110-nifi2.vbrodets.xcu2-8yflx.dev.cldr.work	Running	10.116.106.143	missing data
+0f13c7fa97c97ac7	test110-nifi0.vbrodets.xcu2-8yflx.dev.cldr.work	Running	10.116.105.88	missing data
+08d906275609e1685	test110-nifiscaling0.vbrodets.xcu2-8yflx.dev.cldr.work	Decommissioned	10.116.107.161	missing data
+0863f92a1cc81dc4	test110-nifiscaling1.vbrodets.xcu2-8yflx.dev.cldr.work	Decommissioned	10.116.106.135	missing data

Results

You have now successfully scaled down your NiFi cluster.