

Scaling up or down a NiFi Cluster

Date published: 2019-12-16

Date modified: 2021-10-25

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 three horizontal bars.

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.....	6

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 displays the Cloudera Management Console interface for a NiFi cluster. 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 shows the cluster 'test110' with a 'Running' status and 4 nodes. Below this, there are sections for Environment Details (AWS, vbrodetskiy-aws), Services (CM-UI, NiFi, NiFi Registry, Token Integration), and Cloudera Manager Info (7.4.2 CM version, 7.2.10-1.cdh7.2.10.p0.14376970 runtime version). The Event History section shows recent events such as 'Removed instance(s)' and 'Removing instance(s)'.

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 UI. The top navigation bar includes icons for cluster management and a 'GenerateFlow' button. The number of nodes in the cluster is displayed as '3 / 3', which is highlighted with a red box. The interface also shows various icons for cluster management and a 'GenerateFlow' button.

4. Click Actions, and select Resize.

The Cluster Resize window appears.

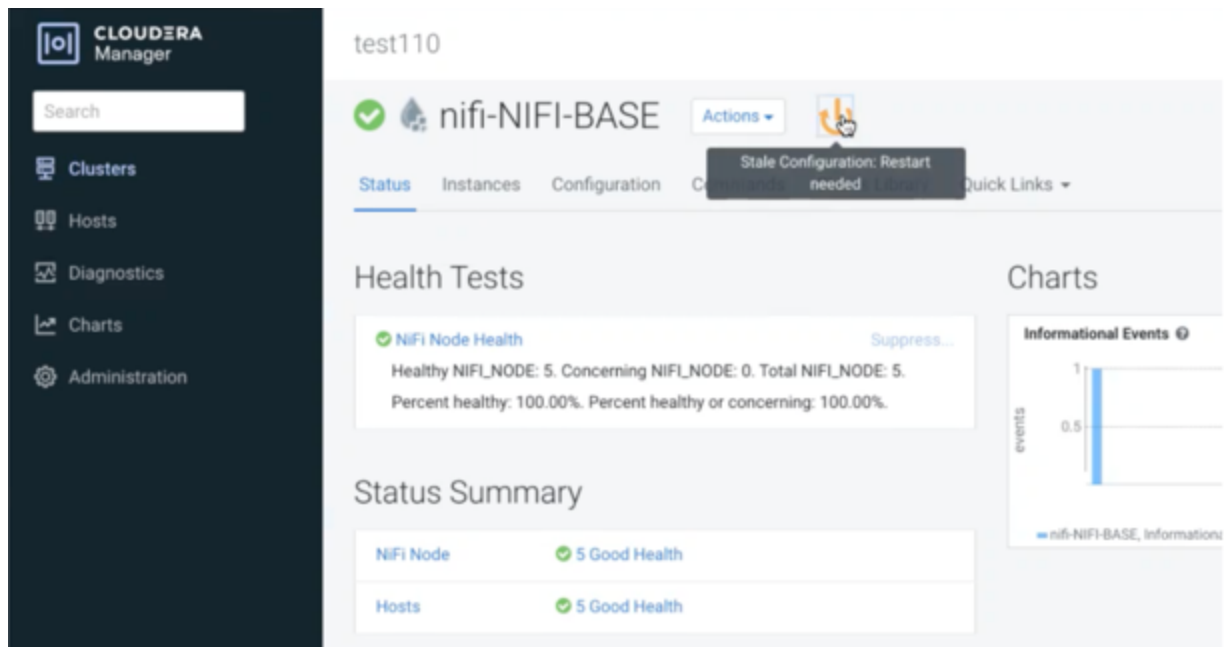
5. Select your hostgroup from the dropdown, increase the number of nodes by clicking the + icon, and then click Resize.



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.

6. Restart the NiFi and NiFi Registry services from the Cloudera Manager UI after you see the new nodes are connected to the NiFi cluster from the NiFi UI.
 - a) Go to the Cloudera Manager UI by clicking the CM-UI link.
The Cloudera Manager UI appears.
 - b) Restart both NiFi and NiFi Registry services.



Results

This step is required to make sure that the legacy nodes are aware of the newly added nodes for the management of permissions and policies. After you restart the services, you successfully scale 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

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:** 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
- 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 URL: https://test110-management0.vbrodets.xcu2-8y8x.dev.cldr.work/test110/cdp-proxy/cm/home/, CM VERSION: 7.4.2, RUNTIME VERSION: 7.2.10-1.cdh7.2.10.p0.14376970, LOGS: Command logs, Service logs
- Management Table:**

ID	FGDN	Status	Private IP	Public IP
+0edd329722b8a4390	test110-management0.vbrodets.xcu2-8y8x.dev.cldr.work	Running	10.116.105.109	missing data

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 shows statistics:

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

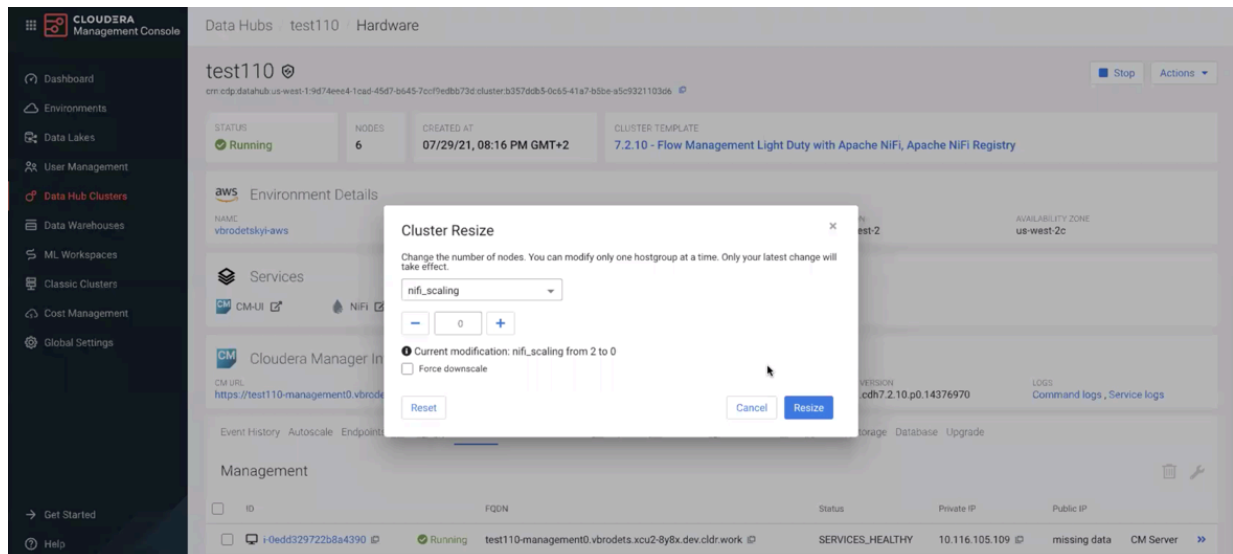
4. Click Actions, and select Resize.

The Cluster Resize window appears.

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



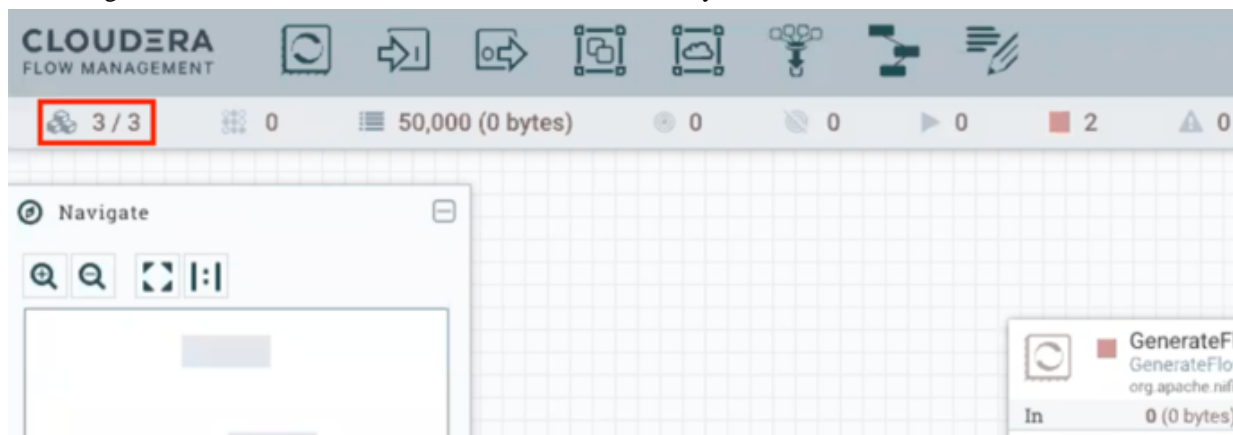
Note: You need to select the hostgroup carefully. For light duty templates, the hostgroup is nifi_scaling. For heavy duty templates, the hostgroup is nifi.



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 Manager UI for a Data Hub cluster. 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'. It includes a 'Cloudera Manager Info' section with version details (7.4.2) and a 'Management' table listing nodes. Below that, there are sections for 'NiFi' and 'NiFi_scaling', each with a table of nodes and their status (Running or Decommissioned).

6. Restart the NiFi and NiFi Registry services from the Cloudera Manager UI after the nodes are removed.

a) Go to the Cloudera Manager UI by clicking the CM-UI link.

The Cloudera Manager UI appears.

b) Restart both NiFi and NiFi Registry services.

The screenshot shows the Cloudera Manager UI for the 'nifi-NIFI-BASE' service. The left sidebar contains navigation options like Clusters, Hosts, Diagnostics, Charts, and Administration. The main content area shows a 'Stale Configuration: Restart needed' warning. Below that, there are sections for 'Health Tests' and 'Status Summary'. The 'Health Tests' section shows 'NiFi Node Health' with a status of 'Healthy NIFI_NODE: 5. Concerning NIFI_NODE: 0. Total NIFI_NODE: 5. Percent healthy: 100.00%. Percent healthy or concerning: 100.00%'. The 'Status Summary' section shows 'NiFi Node' and 'Hosts' both with a status of '5 Good Health'. A 'Charts' section shows 'Informational Events' with a bar chart.

Results

After both services have been restarted, you successfully scale down your NiFi cluster.