Cloudera DataFlow for Data Hub 7.2.14

# Scaling up or down a NiFi Cluster

Date published: 2022-02-24 Date modified: 2022-02-24



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

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.

ELOUDERA Management Console	Data Hubs / test110 / I	Event History			
Dashboard     Environmente	test110 🛞 cm:cdp:datahub:us-west-1:9d74eee4-10	cad-45d7-b645-7ccf9edbb73d:cluster:b357ddb5-0c65-41a	7-b5be-a5c9321103d6		Stop Actions -
C Data Lakes	STATUS N Running 4	NODES         CREATED AT           4         07/29/21, 08:16 PM GMT+2	CLUSTER TEMPLATE 7.2.10 - Flow Management Light	Duty with Apache NiFi, Apache NiFi R	egistry
of Data Hub Clusters 目 Data Warehouses	Environment Deta	DATA LAKE	CREDENTIAL	REGION	AVAILABILITY ZONE
S ML Workspaces 문 Classic Clusters	Data Warehouses     NME     Data Late     CEEDENTIAL     REIGON     AVAA.ABILITY ZONE       ML Workspaces     Image: Clusters     Services     Image: Clusters     Services       Cost Management     Image: Clusters     Image: Clusters     Image: Clusters     Image: Clusters				
Cost Management     Global Settings		iFi 🗗 🎍 NiFi Registry 🗗 📕	Koken Integration		
	Choudera Manage	2FTHTO /brodets.xcu2-8y8x.dev.cldr.work/test110/cdp-pro	cm version ky/cn <sup>2</sup> home/ 7.4.2	RUNTIME VERSION 7.2.10-1.cdh7.2.10.p0.14376970	LOGS Command logs , Service logs
	Event History Autoscale End Events Show All	ipoints (3) Tags (8) Hardware Network Teleme	try Repository Details Image Details Reci	pes (1) Cloud Storage Database Upgrad	e 🛓 DOWNLOAD
→ Get Started ③ Help	<ul> <li>Removed instance(s): i-040 S0/07/2021, 11.43.48</li> <li>Removing instance(s): 2</li> </ul>	c3d714c7b7fc7f5,i-026f180c00067bd08			

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:



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.

Hanagement Console	Data Hubs / test110 /	Event H	listory				
Dashboard     Environments	test110 🛞	1cad-45d7-b6	345-7ccf9edbb73d.cluster.b357ddb5-0c65-41a7-b	5be-a5c9321103d6 🕫			Stop Actions •
El Data Lakes      Ilcer Management	STATUS	NODES	CREATED AT 07/29/21, 08:16 PM GMT+2	CLUSTER TEMPLATE 7.2.10 - Flow Manage	ment Light Duty with A	upache NiFi, Apache NiFi Regi	stry
් Data Hub Clusters	Environment Det	tails			×	N est-2	AVAILABILITY ZONE us-west-2c
더 ML Workspaces 또 Classic Clusters	Services		Cluster Resize Change the number of nodes. You can modify ake effect. niff_scaling	only one hostgroup at a time. O	nly your latest change will		
← Cost Management ④ Global Settings	Cloudera Manag	ier In	2 + O Current modification: nifl_scaling from 0	to 2			
	CM URL https://test110-management0.	vbrode	Reset		Cancel Resize	version .cdh7.2.10.p0.14376970	LOGS Command logs , Service logs
	Event History Autoscale End	dpoints (3)	Tags (8) Hardware Network Telemetry	Repository Details Image I	Details Recipes (1) Clou	d Storage Database Upgrade	
	Events 🛃 Show All	Autosc	ale 📝 Cluster				🛓 DOWNLOAD
→ Get Started	Exercises a standard a version of a standard a ver						
⑦ Help	Removing instance(s): 2						

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

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

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

III CLOUDERA Management Console	Data Hubs / test110	Hardwa	ire			
Dashboard     Environmente	test110 © cm.cdp.datahub:us-west-1:9d74ee	e4-1cad-45d7-b6	45-7ccf9edbb73d:cluster.b357ddb5-0c65-41a7-b	5be-a5c9321103d6 👂		Stop Actions -
Cata Lakes	STATUS	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 NiF	Fi Registry
옷 User Management C <sup>P</sup> Data Hub Clusters	aws Environment I	Details		*		
Data Warehouses     MI Workenacee	NAME vbrodetskyl-aws		DATA LAKE Ø vbrodetskyi-dl	CREDENTIAL eng-dim-daily	REGION us-west-2	AVAILABILITY ZONE us-west-2c
量 Classic Clusters	Services	Nici 🕫	A MC Davister C2	Tokan Internation 178		
<ul> <li>Cost Management</li> <li>Optimized Settings</li> </ul>		NIFI LA		Token Integration		
	CM URL https://test110-management	ager Info	cu2-8y8x.dev.cldr.work/test110/cdp-proxy/	cmf/home/ 7.4.2	RUNTIME VERSION 7.2.10-1.cdh7.2.10.p0.1437697	0 Command logs , Service logs
Cutat Lakes        Cutat Lakes			rade			
	Management					11 F
→ Get Started			FQDN		Status Private	P Public IP
⑦ Help	i-0edd329722b8	a4390 😰	Running test110-management0.v	brodets.xcu2-8y8x.dev.cldr.work	SERVICES_HEALTHY 10.11	6.105.109 D missing data CM Server »

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:

es) 💿 0	0 🥥	▶ 1	1	<b>A</b> 0
		1	0 •	GenerateFlo GenerateFlow org.apache.nifl -
		_	In	0 (0 bytes)
			Read/Write	0 bytes / 0 l
			Out Tasks (Time	20,000 (0 b)
				In Read/Write Out Tasks/Time

4. Click Actions, and select Resize.

The Cluster Resize window appears.

5. 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.

Hanagement Console	Data Hubs / test110 / Hard	ware			
<ul> <li>Dashboard</li> <li>Environments</li> </ul>	test110 🛞 cm.cdp.datahub.us-west-1:9d74eee4-1cad-45d	7-b645-7ccf9edbb73d.cluster.b357ddb5-0c65-41a7-b	Sbe-a5c9321103d6 🖉		Stop Actions -
🕞 Data Lakes	STATUS NODES	CREATED AT 07/29/21, 08:16 PM GMT+2	CLUSTER TEMPLATE 7.2.10 - Flow Management Light Duty w	ith Apache NiFi, Apache NiFi Regi	stry
C <sup>P</sup> Data Hub Clusters	ews Environment Details				
Data Warehouses	vbrodetskyi-aws	Cluster Resize		× est-2	AVAILABILITY ZONE us-west-2c
S ML Workspaces	Services	Change the number of nodes. You can modify take effect.	only one hostgroup at a time. Only your latest change	will	
Global Settings	Cloudera Manager In	O     O     Current modification: nif_scaling from 2     Force downscale	to 0	VESION	
	https://test110-management0.vbrode	Reset	Cancel Resize	.cdh7.2.10.p0.14376970	Command logs , Service logs
	Event History Autoscale Endpoints			torage Database Upgrade	<b>□</b> <i>Y</i>
→ Get Started	- 10	FQDN		Status Private IP	Public IP
(2) Help	📄 📮 i-0edd329722b8a4390 🖻	Running test110-management0.v	brodets.xcu2-8y8x.dev.cldr.work D	SERVICES_HEALTHY 10.116.105.1	09 🗊 missing data CM Server »

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

- a. The nodes to be removed are disconnected from the NiFi cluster.
- b. Processing of data on these nodes is stopped.
- **c.** 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.
- **d.** The NiFi service is stopped on these nodes
- e. 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:

		<u></u> >	់ទ្រាំ		-950 		≡y		
& 3/3 ∰ 0	50,00	0 (0 byte	s)	O	0 Ø	►	0	2	<b>A</b> 0
Ø Navigate	Θ								
ଷ୍ ଷ୍ 🕻 🛛 🛙 🕄 🛛 🕄									
								5	GenerateFlo org.apache.nif
							In		0 (0 bytes)

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:

	Cloudera Manager Info				
	CM VURL CM VURLOW NV VIEW CM V	7.2.10-1.edh7.2.10.p0.1437	6970 Com	i nmand logs , Service logs	
🕞 Data Lakes	Fuerd History Autochain Endoubles (2) Tares (3) Mandaure Notacely Telemetry Denoisies Details Image Details Device	es (1) Cloud Stiviane Database I	Normalie		
Rt, User Management	many county incomes induced in only in comments on and a stand and	er(i) eres erenge somere			
	Management				Ξ,
🖬 Data Warehouses	0 ID FOON	Status Pr	inate (P	Public IP	
5 ML Workspaces	P F0edd829772b8a4390 Ø     P Running text110-management0 vbrodets.xcu2-8y8x.dev.cldr.work Ø	SERVICES_HEALTHY 10	116.105.109 (D	missing data CM Ser	ver
Classic Clusters					
	Nifi				
Global Settings	D io FON	Dates	Private IP	Public IP	
	□ 📮 +0167e5867a0d33cde Ø 🔹 ♦ Furning test110-nif1.vbrodets.xcu2-8y8x.dev.cldr.work Ø	SERVICES_HEALTHY	10.116.107.188 Ø	missing data	
	□	SERVICES_HEALTHY	10.116.106.143 (D	missing data	
	□ 📮 i-0117c7faff7c87ac7 Ø                 ● Running test110-mi60 vbrodets.xou2-8y8x.dev.cldr.work Ø	SERVICES_HEALTHY	10.116.105.88 @	missing data	
	Nifi_scaling				
	0 t t FQN	Status	Private IP	Public IP	
	C I 000000275606e1685 D     A Decommissioned test110 eiffscaling0.vbrodets.xcu2-8y8x.dev.cldr.work	DECOMMISSIONED	10.116.107.161	D missing data	
Pierre Villard	Decommissioned test110 millscaling1 ubrodets you? Bylar devictig work	DECOMMISSIONED	10.116.106.135	D missing data	

#### **Results**

You have now successfully scaled down your NiFi cluster.