Cloudera DataFlow for Data Hub 7.3.1

Scaling up or down a NiFi Cluster

Date published: 2019-12-16 Date modified: 2024-12-10



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	

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 / E	Event History			
 Dashboard Environments 	test110 🛞 cm:cdp:datahub:us-west-1:9d74eee4-1c	cad-45d7-b645-7ccf9edbb73d.cluster.b357ddb5-0c65-41a7	7-b5be-a5c9321103d6		Stop Actions •
🗟: Data Lakes	STATUS N Running 4	KODES 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 Re	egistry
옷 User Management ♂ Data Hub Clusters	Environment Deta	ails DATA LAKE ✔ vbrodetskvi-di	CREDENTIAL eng-dim-daily	REGION us-west-2	AVAILABILITY ZONE
与 ML Workspaces 曜 Classic Clusters	Services	C References	engourrouny	us nest z	ue moar zu
্যে Cost Management 🞯 Global Settings	_		Token Integration 🖪		
	Cloudera Manage	er Info brodets.xcu2-8y8x.dev.cldr.work/test110/cdp-prox	cm version ay/cn home/ 7.4.2	RUNTIME VERSION 7.2.10-1.cdh7.2.10.p0.14376970	LOGS Command logs , Service logs
		points (3) Tags (8) Hardware Network Telemer	try Repository Details Image Details Reci	pes (1) Cloud Storage Database Upgrade	e
→ Get Started ③ Help	 Removed instance(s): i-040 30/07/2021,1143.48 Removing instance(s): 2 	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 I	History				
 Dashboard Environments 	test110 @ cm:cdp.datahub:us-west-1:9d74eee4	-1cad-45d7-	o645-7ccf9edbb73d:cluster:b357ddb5-0c65-41a7-b	556-a5c9321103d6 👂			Stop Actions -
🗟 Data Lakes	STATUS	NODES 4	CREATED AT 07/29/21, 08:16 PM GMT+2	CLUSTER TEMPLATE 7.2.10 - Flow Manageme	ent Light Duty with A	pache NiFi, Apache NiFi Regis	stry
옷 User Management c ^o Data Hub Clusters	AWS Environment De	tails		00000000		Nest-2	AVAILABILITY ZONE
등 ML Workspaces 뗥 Classic Clusters	Services	1	Cluster Resize Change the number of nodes. You can modify take effect.	only one hostgroup at a time. Only	× your latest change will		
Cost Management Global Settings	🖼 см-и 🖻 🌒 і	nifi 🖪	iff_scaling				
	CM URL https://test110-management0		Current modification: nin_scaling from C		ancel Resize	version .cdh7.2.10.p0.14376970	LOGS Command logs , Service logs
) Tags (8) Hardware Network Telemetry scale 🛃 Cluster	Repository Details Image Det	ails Recipes (1) Clou	d Storage Database Upgrade	🛓 DOWNLOAD
→ Get Started ⑦ Help	 Removed instance(s): i-0 30/07/2021, 11.43.48 Removing instance(s): 2 	4c3d714c	7b7fc7f5,i-026f180c00067bd08				

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.

CLOUDERA Management Console	Data Hubs / test110 / Hardw	vare			
Dashboard Environments	test110 🛞 cm:cdp:datahub:us-west-1:9d74eee4-1cad-45d7	b645-7ccf9edbb73d:cluster.b357ddb5-0c65-41a7-	55be-a5c9321103d6 🙎		Stop Actions •
C Data Lakes	STATUS NODES	CREATED AT 07/29/21, 08:16 PM GMT+2	CLUSTER TEMPLATE 7.2.10 - Flow Management Light Du	ıty with Apache NiFi, Apache NiFi F	Registry
옷 User Management 야 Data Hub Clusters	ews Environment Details		*		
Data Warehouses	NAME vbrodetskyi-aws	DATA LAKE Øvbrodetskyi-di	CREDENTIAL eng-dim-daily	REGION us-west-2	AVAILABILITY ZONE us-west-2c
S ML Workspaces 臣 Classic Clusters	Services				
Cost Management	🖾 cm-ui 🗹 🍐 Nifi 🕻	🖕 NiFi Registry 🗗 🔣	Token Integration 🖪		
Ø Global Settings		xcu2-8y8x.dev.cldr.work/test110/cdp-proxy	/cmf/home/ 7.4.2	RUNTIME VERSION 7.2.10-1.cdh7.2.10.p0.14376970 s (1) Cloud Storage Database Upgrad	LOOS Command logs , Service logs
	Management				Ĩ <i>F</i>
 → Get Started ⑦ Help 	 ID iD i-0edd329722b8a4390 iD 	FQDN © Running test110-management0.	vbrodets.xcu2-8y8x.dev.cldr.work D	Status Private IP SERVICES_HEALTHY 10.116.1	Public IP 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:

	↓				1	
& 5/5 ∰ 0	🔳 50,000 (0 byt	es) 💿 0	0 🧭	▶ 1	1	A 0
Ø Navigate	Θ					
ଷ୍ ସ୍ 🕻 🛙 🛙						
					0 •	GenerateFlow org.apache.nifl ·
					In	0 (0 bytes)
					Read/Write	
					Out	20,000 (0 b)
L					Tasks/Tim	e 20,000/00

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			
 Dashboard Environments 	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
옷 User Management C ^P Data Hub Clusters	ews Environment Details				
Data Warehouses	NAME vbrodetskyi-aws	Cluster Resize		× est-2	avaiLabiLity zone us-west-2c
5 ML Workspaces 曼 Classic Clusters 슈 Cost Management	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	VERSION	
	https://test110-management0.vbrode	Reset	Cancel Resize	e.cdh7.2.10.p0.14376970	Command logs , Service logs
	Event History Autoscale Endpoints			lorage Database Upgrade	□ <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:

				- <i>V</i>		
8 € 3/3 ∰ 0	I≣ 50,000 (0 byt	es) 💿 0	0 🚿	▶ 0	2	A 0
Ø Navigate						
ଷ୍ୟ 🕻 ା∶ା						
					0	GenerateFlo org.apache.nifi
					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:

	u2-Bylix.dev.cidx.work/text110/cdp.proxy.cmf/horne/ Tags (8) Handware Network Telerretry Repository.Details			Command logs , Service logs
	Tags (8) Handware Network Telemetry Repository Details			
			toriage Database Upgrade	
Management				
management				
0.0	FODM	Status	Private IP	Public IP
🔲 🖵 i-0edd32972258a4390 Ø	Running test110-management0.vbrodets.xcu2-8y8x.	iev.cldr.work (D SERVICE	S_HEALTHY 10.116.105.109 (D	missing data CM Serv
Nifi				
0 0	FION	Status	Private IP	Public IP
D +0107e5867a0d33cde Ø	Running test110-nifi1.vbrodets.xcu2-8y8x.dev	ids.work (D SERVICE	S_HEALTHY 10.116.107.188	Ø missing data
Image: 107a7a3c02a540ax2c iP	Running test110-nifi2.vbrodets.xcu2-8y8x.dev.	Idc.work (2 SERVICE	S_HEALTHY 10.116.106.143	D missing data
	Running test110-nifi0.vbrodets.xcu2-8y8x.dev	Ide work @ SERVICE	S_HEALTHY 10.116.105.88 G	D missing data
Nifi_scaling				
0	† FQM		Status Private IP	Public IP
	P +0ed502972258e4390 € Nifi i		Image: Interest in the second secon	Image: Constraint of the second sec

Results

You have now successfully scaled down your NiFi cluster.