

Kết quả báo cáo nhóm RHUST-04

- Quá trình cài đặt thành công.
- Bọn em cài đặt Hadoop trên hệ điều hành ubuntu 20.04
- Cụm sử dụng 3 máy trong đó có 1 máy là nameNode và 2 máy là dataNode.

hình ảnh minh chứng .

The screenshot displays the Hadoop cluster management interface in a web browser. The page title is "Nodes of the cluster". The interface includes a sidebar with navigation options like "Cluster", "About", "Nodes", "Node Labels", and "Applications". The main content area shows "Cluster Metrics" and "Scheduler Metrics". Below these, there is a table listing the nodes in the cluster.

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Mem Used	Mem Avail	VCores Used	VCores Avail	Version
/default-rack	/default-rack	RUNNING	slave-2-36195	slave-2-8042	Sat Oct 17 00:19:59 +0700 2020		0	0 B	8 GB	0	8	2.7.7
/default-rack	/default-rack	RUNNING	slave-1-33285	slave-1-8042	Sat Oct 17 00:19:57 +0700 2020		0	0 B	8 GB	0	8	2.7.7

Showing 1 to 2 of 2 entries

```

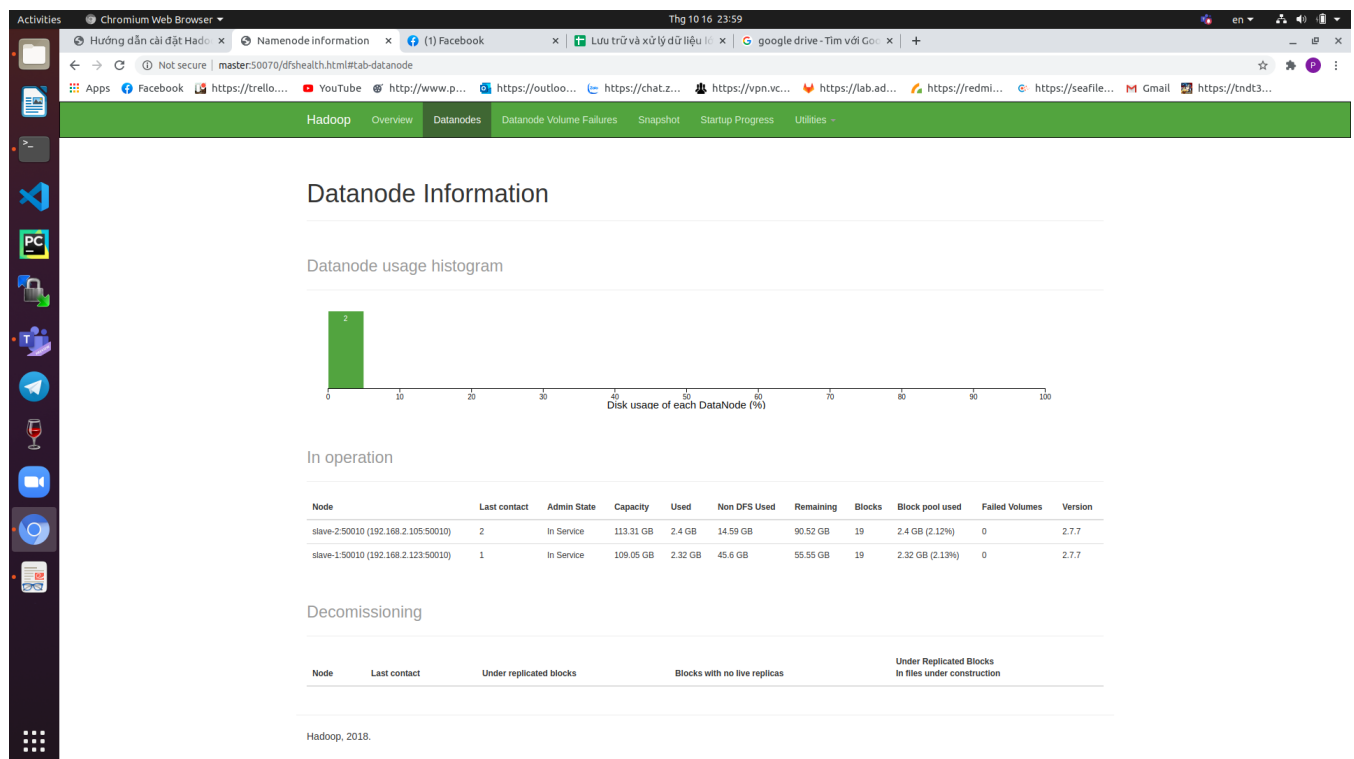
hadoopuser@master:~$ hdfs dfsadmin -report
20/10/16 23:55:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Configured Capacity: 238753411072 (222.36 GB)
Present Capacity: 161913181937 (150.79 GB)
DFS Remaining: 156844466176 (146.07 GB)
DFS Used: 5088715761 (4.72 GB)
DFS Used%: 3.13%
Under replicated blocks: 19
Blocks with corrupt replicas: 0
Missing blocks: 0
Missing blocks (with replication factor 1): 0

-----
Live datanodes (2):

Name: 192.168.2.105:50010 (slave-2)
Hostname: slave-2
Decommission Status : Normal
Configured Capacity: 121665052672 (113.31 GB)
DFS Used: 2574734197 (2.40 GB)
Non DFS Used: 15665347723 (14.59 GB)
DFS Remaining: 97200676864 (90.53 GB)
DFS Used%: 2.12%
DFS Remaining%: 79.89%
Configured Cache capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xcelsers: 1
Last contact: Fri Oct 16 23:55:44 ICT 2020

Name: 192.168.2.123:50010 (slave-1)
Hostname: slave-1
Decommission Status : Normal
Configured Capacity: 117088358400 (109.05 GB)
DFS Used: 2493981564 (2.32 GB)
Non DFS Used: 48958712964 (45.60 GB)
DFS Remaining: 59643789312 (55.55 GB)
DFS Used%: 2.13%
DFS Remaining%: 50.94%
Configured Cache capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xcelsers: 1
Last contact: Fri Oct 16 23:55:45 ICT 2020

```



- Bạn em lưu trữ file 1.zip có kích thước là 2,28GB vs lệnh :
 - hdfs dfs -D dfs.replication=3 -copyFromLocal /home/phucnq/PycharmProjects/PycharmProjects/1.zip /phucnq
 - ở đây file được tạo bản sao và dữ liệu được phân chia ra 2 dataNode.

Activities Chromium Web Browser Thg 10 16 23:59

Hướng dẫn cài đặt Hadoop x Browsing HDFS x (1) Facebook x Lưu trữ và xử lý dữ liệu l... x google drive - Tìm với Go... x +

Not secure | master50070/explorer.html#/phucnq

Apps Facebook https://trello... YouTube http://www.p... https://outloo... https://chat.z... https://vpn.vc... https://lab.ad... https://redmi... https://seaf... Gmail https://tndt3...

Hadoop Overview Datanodes Snapshot Startup Progress Utilities -

Browse Directory

/phucnq

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hadoopuser	supergroup	2.28 GB	10/16/2020, 11:50:06 PM	3	128 MB	1.zip

Hadoop, 2018.

-Dữ liệu mặc định với mặc định với mỗi block là 128MB nên do đó dữ liệu được chia thành 19 block.