

Big Data Analytics – Lab Project

Number of machines used:- 3 (1 Master, 2 Slaves)

Master and Slaves IP Addresses

Master – 10.0.0.4

Slave 1 – 10.0.0.5

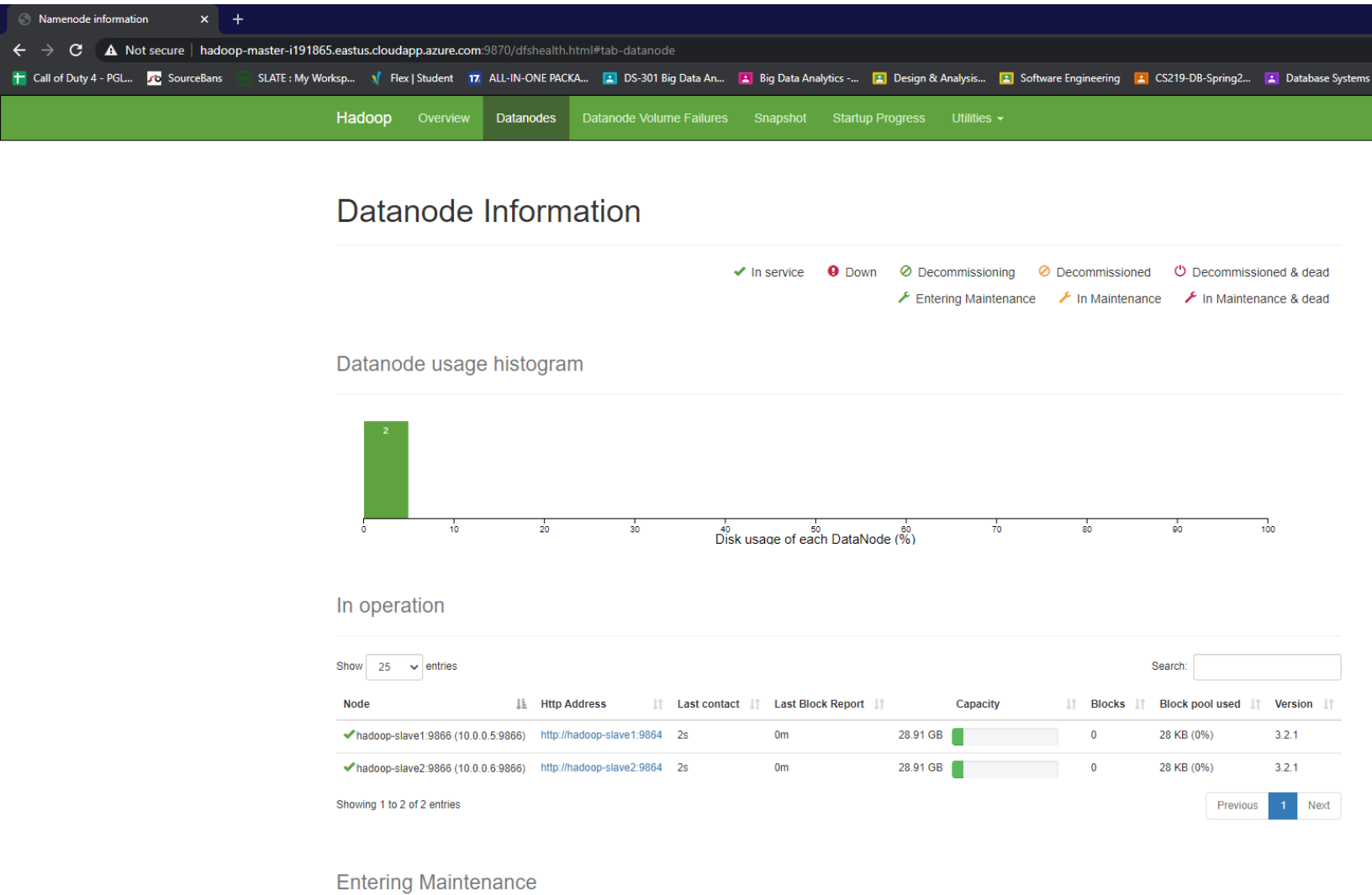
Slave 2 – 10.0.0.6

```
GNU nano 2.9.3 /etc/hosts
27.0.0.1 localhost
10.0.0.4 hadoop-master
10.0.0.5 hadoop-slave1
10.0.0.6 hadoop-slave2
# The following lines are desirable for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
ff02::3 ip6-allhosts
```

I

```
^G Get Help      ^O Write Out    ^W Where Is     ^K Cut Text     ^J Justify      ^C Cur Pos      ^- Undo         ^-A Mark Text    ^-] To Bracket  ^-A Previous    ^B Back         ^-E Exit         ^R Read File    ^M Replace      ^U Uncut Text   ^T To Spell     ^_ Go To Line    ^-R Redo        ^-F Copy Text   ^-W WhereIs Next ^-V Next        ^E Forward
```

Data Nodes (Screenshot 1)



Node States (Screenshot 2)

Nodes of the cluster

Not secure | [hadoop-master-i191865.eastus.cloudapp.azure.com:3088/cluster/nodes](#)

hadoop

Nodes of the cluster

Cluster

About

Nodes

Node Labels

Applications

NEW

NEW_SAVING

SUBMITTED

ACCEPTED

RUNNING

FINISHED

FAILED

KILLED

Scheduler

Tools

Cluster Metrics

1	Apps Submitted	0	Apps Pending	0	Apps Running	1	Apps Completed	0	Containers Running	0 B	Memory Used	16 GB	Memory Total	0 B	Memory Reserved	0	VCores Used	16	VCores Total	0	VCores Reserved
---	----------------	---	--------------	---	--------------	---	----------------	---	--------------------	-----	-------------	-------	--------------	-----	-----------------	---	-------------	----	--------------	---	-----------------

Cluster Nodes Metrics

2	Active Nodes	0	Decommissioning Nodes	0	Decommissioned Nodes	0	Lost Nodes	0	Unhealthy Nodes	0	Rebooted Nodes	0	Shutdown Nodes
---	--------------	---	-----------------------	---	----------------------	---	------------	---	-----------------	---	----------------	---	----------------

Scheduler Metrics

Capacity Scheduler	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum Cluster Application Priority
	[memory-mb (unit=Mi), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>	0

Showing 1 to 2 of 2 entries

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Allocation Tags	Mem Used	Mem Avail	VCores Used	VCores Avail	Version
/default-rack		RUNNING	hadoop-slave1:40795	hadoop-slave1:8042	Tue May 11 13:26:23 +0000 2021		0		0 B	8 GB	0	8	3.2.1
/default-rack		RUNNING	hadoop-slave2:41335	hadoop-slave2:8042	Tue May 11 13:26:23 +0000 2021		0		0 B	8 GB	0	8	3.2.1

Showing 1 to 2 of 2 entries

Success – Output (Screenshot 3)




Browsing HDFS

Not secure | [hadoop-master-i191865.eastus.cloudapp.azure.com:9870/explorer.html#/output/i191865](#)



Call of Duty 4 - PGL... SourceBans SLATE : My Worksp... Flex | Student ALL-IN-ONE PACKA... Global Order Tracki... DS-301 Big Data An... Big Data Analytics -... Design & Analysis... Software Engineering CS219-DB-Spring2...

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Browse Directory

/output/i191865 Go!   

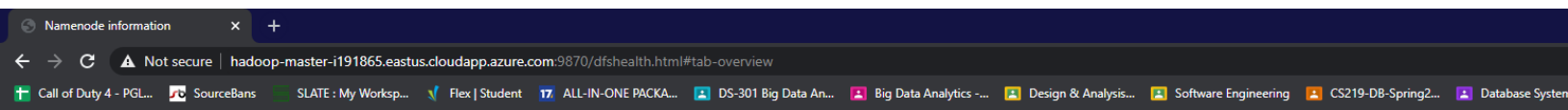
Show entries Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	-rw-r--r--	hadoopuser	supergroup	0 B	May 26 22:24	2	128 MB	_SUCCESS	
<input type="checkbox"/>	-rw-r--r--	hadoopuser	supergroup	13.38 MB	May 26 22:24	2	128 MB	part-00000	

Showing 1 to 2 of 2 entries Previous **1** Next

Hadoop, 2019.

Cluster Overview (Screenshot 4)



Summary

Security is off.

Safemode is off.

1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).

Heap Memory used 74.98 MB of 313.5 MB Heap Memory. Max Heap Memory is 1.73 GB.

Non Heap Memory used 55.73 MB of 56.84 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	57.81 GB
Configured Remote Capacity:	0 B
DFS Used:	56 KB (0%)
Non DFS Used:	6.12 GB
DFS Remaining:	51.66 GB (89.36%)
Block Pool Used:	56 KB (0%)
DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 0.00% / 0.00%
Live Nodes	2 (Decommissioned: 0, In Maintenance: 0)
Dead Nodes	0 (Decommissioned: 0, In Maintenance: 0)
Decommissioning Nodes	0
Entering Maintenance Nodes	0
Total Datanode Volume Failures	0 (0 B)
Number of Under-Replicated Blocks	0
Number of Blocks Pending Deletion (including replicas)	0
Block Deletion Start Time	Tue May 11 18:22:09 +0500 2021
Last Checkpoint Time	Tue May 11 18:00:52 +0500 2021
Enabled Erasure Coding Policies	RS-6-3-1024k

View Output (Screenshot 5)

The screenshot shows a web browser window with the URL `hadoop-master-i191865.eastus.cloudapp.azure.com:9870/explorer.html#/output/i191865`. The page title is "Browse Directory". A message "Couldn't preview the file." is displayed. Below it, the path `/output/i191865` is shown. A table lists file entries with columns for "Permission" and "Owner". The table shows two entries, both with permissions `-rw-r--r--` and owner `hadoopu`. Below the table, it says "Showing 1 to 2 of 2 entries". At the bottom, it says "Hadoop, 2019."

A modal window titled "File information - part-00000" is open. It has three tabs: "Download", "Head the file (first 32K)", and "Tail the file (last 32K)". The "Block information" section shows a dropdown menu set to "Block 0". The details listed are:

- Block ID: 1073741844
- Block Pool ID: BP-1418483291-10.0.0.4-1620738052145
- Generation Stamp: 1020
- Size: 14029448
- Availability:
 - hadoop-slave2
 - hadoop-slave1

The "File contents" section shows a text area with the placeholder text "[object Object]". A "Close" button is located at the bottom right of the modal.

Cat Output (Screenshot 6)

hadoopuser@hadoop-master: ~/data

```
~serenashay.com~ 1
~sharon 2
~sherry 2
~shrugs~ 1
~smashwords 4
~smashwords.com/books/view/391592~ 1
~solaragordon.wordpress.com~ 1
~song 1
~state 3
~steven 1
~summer 4
~s~ 26
~the 34
~titlemagic.blogspot.com~ 1
~transcendental 1
~trystan~ 74
~tyrstan~ 2
~weathercock 2
~what 1
~whoops~ 2
~wild 1
~will 4
~william 2
~will~ 7
~winter 2
~wood 1
~word 1
~wunder~ 1
~you 2
~your 3
~ziah~ 2
~z~ 302
~z~ 2
~z~ 104
~book 1
~i 1
~instead 1
~my 1
~other 2
~tearing 1
~z~ 1722
~z~ 2
~z~ 33
~z~ 18
~z~ 320
~z~ 8
~z~ 34
~z~ 315
~z~ 45
~z~ 12
~z~ 1
~z~ 15
~z~ 157
~z~ 1
~z~ 1
~z~ 3
~z~ 3
~z~ couple 1
~z~ 2
~z~ 1
~z~ 1
hadoopuser@hadoop-master:~/data$
```

VM View (Screenshot 7)

Virtual machines - Microsoft Azure

portal.azure.com/#blade/HubsExtension/BrowseResource/resourceType/Microsoft.Compute%2FVirtualMachines

Microsoft Azure

Search resources, services, and docs (G+/I)

FAST NATIONAL UNIVERSITY (nu.edu.pk)

Home >

Virtual machines

FAST National University (nu.edu.pk)

+ Add

Switch to classic

Reservations

Manage view

Refresh

Export to CSV

Open query

Feedback

Leave preview

Assign tags

Start

Restart

Stop

Delete

Services

Maintenance

Filter for any field...

Subscription == all

Resource group == all

Location == all

Add filter

Showing 1 to 3 of 3 records.

No grouping

List view

<input type="checkbox"/>	Name ↑↓	Subscription ↑↓	Resource group ↑↓	Location ↑↓	Status ↑↓	Operating system ↑↓	Size ↑↓	Public IP address ↑↓	Disks ↑↓	
<input type="checkbox"/>	hadoop-master	Azure for Students	Hadoop	East US	Stopped (deallocated)	Linux	Standard_D2s_v3	52.224.160.64	1	...
<input type="checkbox"/>	hadoop-slave1	Azure for Students	Hadoop	East US	Stopped (deallocated)	Linux	Standard_D2s_v3	40.71.223.184	1	...
<input type="checkbox"/>	hadoop-slave2	Azure for Students	Hadoop	East US	Stopped (deallocated)	Linux	Standard_E2s_v3	40.87.89.125	1	...

< Previous

Page 1 of 1

Next >