

Assignment -1

1. Start Hadoop single node on Acadgild VM. The command is **start-all.sh**

Answer 1:

```
[acadgild@localhost ~]$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
18/02/11 14:28:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-namenode-localhost.localdomain.out
localhost: starting datanode, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-datanode-localhost.localdomain.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-secondarynamenode-localhost.localdomain.out
18/02/11 14:28:27 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
starting yarn daemons
starting resourcemanager, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-resourcemanager-localhost.localdomain.out
localhost: starting nodemanager, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-nodemanager-localhost.localdomain.out
[acadgild@localhost ~]$
```

This command is used to start all the daemons in Hadoop.

2. Run JPS command to see if all Hadoop daemons are running.

Answer 2:

```
[acadgild@localhost ~]$ jps
4224 DataNode
4530 ResourceManager
4387 SecondaryNameNode
4099 NameNode
4632 NodeManager
4959 Jps
[acadgild@localhost ~]$
```

JPS command is used to check all the Hadoop daemons like NameNode, DataNode, ResourceManager, NodeManager etc. which are running on the machine.

3. Run few unix commands like pwd, ls -ls, etc.

Answer 3:

pwd command: Gets current working directory.

```
[acadgild@localhost ~]$ pwd
/home/acadgild
```

ls -ls command: Lists all the set of directories in the current path.

```
[acadgild@localhost ~]$ ls -ls
total 56
4 -rw-rw-r--. 1 acadgild acadgild 107 Feb 10 13:39 abc.txt
4 drwxr-xr-x. 3 acadgild acadgild 4096 Feb  2 12:51 Desktop
4 drwxr-xr-x. 2 acadgild acadgild 4096 Feb  2 12:52 Documents
4 drwxr-xr-x. 2 acadgild acadgild 4096 Feb  2 12:51 Downloads
4 drwxrwxr-x. 3 acadgild acadgild 4096 Dec 29 16:59 eclipse
4 drwxrwxr-x. 3 acadgild acadgild 4096 Jan 16 14:02 eclipse-workspace
4 drwxrwxr-x. 12 acadgild acadgild 4096 Feb  2 12:44 install
4 drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Music
4 drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Pictures
4 drwxrwxr-x. 2 acadgild acadgild 4096 Feb 10 17:31 Practice
4 drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Public
4 drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Templates
4 -rw-rw-r--. 1 acadgild acadgild  52 Feb 10 11:32 test.txt
4 drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Videos
[acadgild@localhost ~]$
```

mkdir command: To make a new directory.

Refer the yellow highlighted directory in the below screenshot.

```
[acadgild@localhost ~]$ mkdir Hadoop-Practice
[acadgild@localhost ~]$ ls -l
total 60
-rw-rw-r--. 1 acadgild acadgild 107 Feb 10 13:39 abc.txt
drwxr-xr-x. 3 acadgild acadgild 4096 Feb  2 12:51 Desktop
drwxr-xr-x. 2 acadgild acadgild 4096 Feb  2 12:52 Documents
drwxr-xr-x. 2 acadgild acadgild 4096 Feb  2 12:51 Downloads
drwxrwxr-x. 3 acadgild acadgild 4096 Dec 29 16:59 eclipse
drwxrwxr-x. 3 acadgild acadgild 4096 Jan 16 14:02 eclipse-workspace
drwxrwxr-x. 2 acadgild acadgild 4096 Feb 11 14:44 Hadoop-Practice
drwxrwxr-x. 12 acadgild acadgild 4096 Feb  2 12:44 install
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Music
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Pictures
drwxrwxr-x. 2 acadgild acadgild 4096 Feb 10 17:31 Practice
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Public
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Templates
-rw-rw-r--. 1 acadgild acadgild  52 Feb 10 11:32 test.txt
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Videos
[acadgild@localhost ~]$
```

Cat command: To see the content of the file.

```
[acadgild@localhost ~]$ cat abc.txt
Welcome to Hadoop. This is Ankit Murali.
A mainframe developer ---> transforming to be Big Data analyst!
```

Copy command: To copy the file from one to another file use command “cp <From_file> <To_file>”

Listing is done here to check the new file in the list and again “cat” command is used to see the content.

```
[acadgild@localhost ~]$ cp abc.txt xyz.txt
[acadgild@localhost ~]$ ls -l
total 64
-rw-rw-r--. 1 acadgild acadgild 107 Feb 10 13:39 abc.txt
drwxr-xr-x. 3 acadgild acadgild 4096 Feb  2 12:51 Desktop
drwxr-xr-x. 2 acadgild acadgild 4096 Feb  2 12:52 Documents
drwxr-xr-x. 2 acadgild acadgild 4096 Feb  2 12:51 Downloads
drwxrwxr-x. 3 acadgild acadgild 4096 Dec 29 16:59 eclipse
drwxrwxr-x. 3 acadgild acadgild 4096 Jan 16 14:02 eclipse-workspace
drwxrwxr-x. 2 acadgild acadgild 4096 Feb 11 14:44 Hadoop-Practice
-rw-rw-r--. 1 acadgild acadgild  77 Feb 11 15:04 Hadoop_Practice.txt
drwxrwxr-x. 12 acadgild acadgild 4096 Feb  2 12:44 install
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Music
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Pictures
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Public
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Templates
-rw-rw-r--. 1 acadgild acadgild  52 Feb 10 11:32 test.txt
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Videos
-rw-rw-r--. 1 acadgild acadgild 107 Feb 16 22:50 xyz.txt
[acadgild@localhost ~]$ cat xyz.txt
Welcome to Hadoop. This is Ankit Murali.
A mainframe developer ---> transforming to be Big Data analyst!
[acadgild@localhost ~]$
```

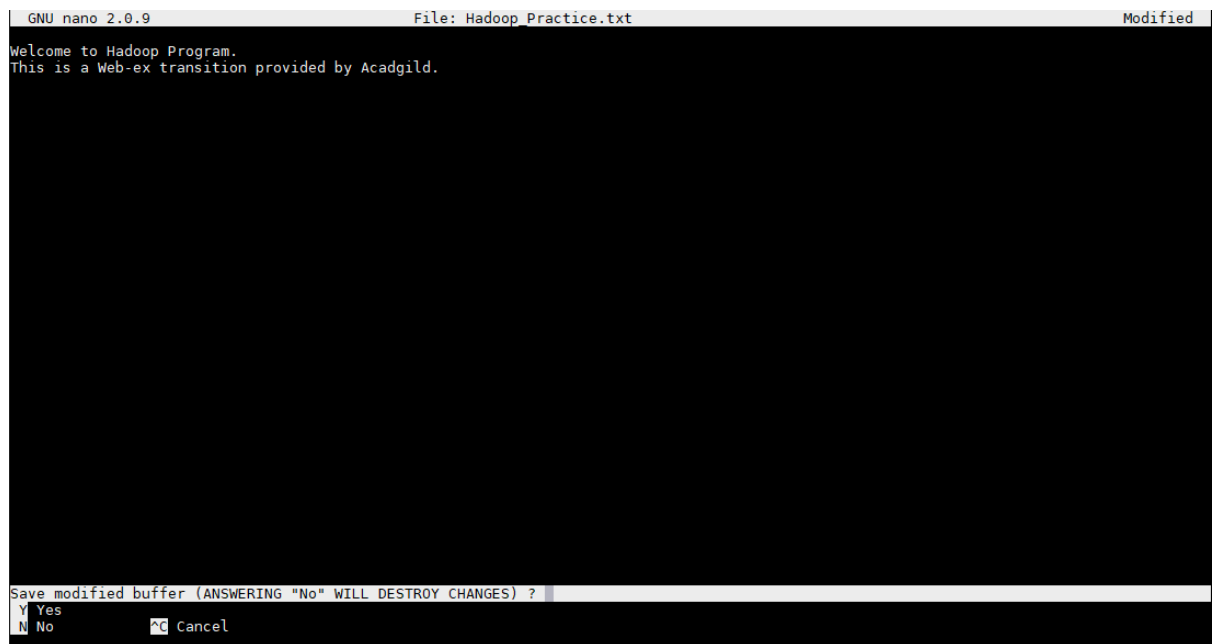
4. Create a file from the terminal using nano editor (example: nano test.txt), add some content in it. Cat it to see if the content is saved.

Answer 4:

Use the below command to create file in terminal.

```
[acadgild@localhost ~]$ nano Hadoop_Practice.txt
[acadgild@localhost ~]$
```

This command will open space to enter the data in Hadoop_Practice.txt file as in below screenshot.



Press Y to save and then press Enter to return to terminal

Use cat command to see the content of file - **Hadoop_Practice.txt** is saved.

```
[acadgild@localhost ~]$ cat Hadoop_Practice.txt
Welcome to Hadoop Program.
This is a Web-ex transition provided by Acadgild.
[acadgild@localhost ~]$
```

5. Open the hdfs web page by typing *localhost: 50070* in the browser. Check all the details of the HDFS.

Answer 5:

The above link opens all the details of HDFS in the web browser only if all the daemons are started in Hadoop.

If the daemons are not started then the web page won't be opened.

Applications Places System Sun Feb 11, 9:51 PM Acadgild

Namenode information - Mozilla Firefox

Namenode information x +

localhost:50070/dfshealth.html#tab-overview

Overview 'localhost:8020' (active)

Started:	Sun Feb 11 21:48:36 IST 2018
Version:	2.6.5, re8c9fe0b4c252caf2ebf1464220599650f119997
Compiled:	2016-10-02T23:43Z by sjlee from branch-2.6.5
Cluster ID:	CID-7b3f9bd8-f34c-4fb8-87aa-f76b6dfbd809
Block Pool ID:	BP-437583619-127.0.0.1-1517555661954

Summary

Security is off.

Safemode is off.

27 files and directories, 10 blocks = 37 total filesystem object(s).

Heap Memory used 42.43 MB of 60.06 MB Heap Memory. Max Heap Memory is 966.69 MB.

Non Heap Memory used 44.07 MB of 45.06 MB Committed Non Heap Memory. Max Non Heap Memory is -1 B.

Configured Capacity:	17.11 GB
DFS Used:	282.84 MB

Namenode informatio... acadgild@localhost:~

Applications Places System Sun Feb 11, 9:51 PM Acadgild

Namenode information - Mozilla Firefox

Namenode information x +

localhost:50070/dfshealth.html#tab-overview

Non Heap Memory used 44.07 MB of 45.06 MB Committed Non Heap Memory. Max Non Heap Memory is -1 B.

Configured Capacity:	17.11 GB
DFS Used:	282.84 MB
Non DFS Used:	12.23 GB
DFS Remaining:	4.61 GB
DFS Used%:	1.61%
DFS Remaining%:	26.92%
Block Pool Used:	282.84 MB
Block Pool Used%:	1.61%
DataNodes usages% (Min/Median/Max/stdDev):	1.61% / 1.61% / 1.61% / 0.00%
Live Nodes	1 (Decommissioned: 0)
Dead Nodes	0 (Decommissioned: 0)
Decommissioning Nodes	0
Number of Under-Replicated Blocks	3
Number of Blocks Pending Deletion	0
Block Deletion Start Time	2/11/2018, 9:48:36 PM

Namenode informatio... acadgild@localhost:~

Applications Places System Sun Feb 11, 9:52 PM Acadgild

Namenode information - Mozilla Firefox

Namenode information x +

localhost:50070/dfshealth.html#tab-overview Search

Number of Blocks Pending Deletion	0
Block Deletion Start Time	2/11/2018, 9:48:36 PM

NameNode Journal Status

Current transaction ID: 482

Journal Manager	State
FileJournalManager(root=/home/acadgild/install/data/dfs/name)	EditLogFileOutputStream(/home/acadgild/install/data/dfs/name/current/edits_inprogress_000000000000000000482)

NameNode Storage

Storage Directory	Type	State
/home/acadgild/install/data/dfs/name	IMAGE_AND_EDITS	Active

Hadoop, 2016. Legacy UI

Namenode informatio... acadgild@localhost:~

Data Nodes information:

Applications Places System Sun Feb 11, 9:53 PM Acadgild

Namenode information - Mozilla Firefox

Namenode information x +

localhost:50070/dfshealth.html#tab-datanode

Hadoop Overview Datanodes Snapshot Startup Progress Utilities

Datanode Information

In operation

Node	Last contact	Admin State	Capacity	Used	Non DFS Used	Remaining	Blocks	Block pool used	Failed Volumes	Version
localhost (127.0.0.1:50010)	2	In Service	17.11 GB	282.84 MB	12.23 GB	4.61 GB	10	282.84 MB (1.61%)	0	2.6.5

Decommissioning

Node	Last contact	Under replicated blocks	Blocks with no live replicas	Under Replicated Blocks In files under construction
------	--------------	-------------------------	------------------------------	---

Namenode informatio... acadgild@localhost:~

Applications Places System Sun Feb 11, 9:54 PM Acadgild

Namenode information - Mozilla Firefox

Namenode information x +

localhost:50070/dfshealth.html#tab-snapshot

Hadoop Overview Datanodes Snapshot Startup Progress Utilities

Snapshot Summary

Snapshottable directories: 0

Path	Snapshot Number	Snapshot Quota	Modification Time	Permission	Owner	Group
------	-----------------	----------------	-------------------	------------	-------	-------

Snapshotted directories: 0

Snapshot ID	Snapshot Directory	Modification Time
-------------	--------------------	-------------------

Hadoop, 2016. Legacy UI

Namenode informatio... acadgild@localhost:~

ApplicationsPlacesSystem

Sun Feb 11, 9:55 PMAcadgild

Namenode information - Mozilla Firefox

Namenode information

localhost:50070/dfshealth.html#tab-startup-progress

Search

Hadoop

Overview

Datanodes

Snapshot

Startup Progress

Utilities

Startup Progress

Elapsed Time: 38 sec, Percent Complete: 100%

Phase	Completion	Elapsed Time
Loading fsimage /home/acadgild/install/data/dfs/name/current/fsimage_00000000000000000479 2.16 KB	100%	0 sec
inodes (0/0)	100%	
delegation tokens (0/0)	100%	
cache pools (0/0)	100%	
Loading edits	100%	0 sec
/home/acadgild/install/data/dfs/name/current/edits_00000000000000000480-00000000000000000480 1 MB (1/1)	100%	
/home/acadgild/install/data/dfs/name/current/edits_00000000000000000481-00000000000000000481 1 MB (1/1)	100%	
Saving checkpoint	100%	0 sec

Namenode informatio...

acadgild@localhost: ~

The screenshot displays the Hadoop Namenode information page in a Mozilla Firefox browser. The page title is "Namenode information - Mozilla Firefox". The address bar shows the URL "localhost:50070/dfshealth.html#tab-startup-progress". The page content is organized into sections, each with a progress bar and a status indicator.

Section	Progress	Status
delegation tokens (0/0)	100%	
cache pools (0/0)	100%	
Loading edits	100%	0 sec
/home/acadgild/install/data/dfs/name/current /edits_00000000000000000480-00000000000000000480 1 MB (1/1)		
/home/acadgild/install/data/dfs/name/current /edits_00000000000000000481-00000000000000000481 1 MB (1/1)		
Saving checkpoint	100%	0 sec
inodes /home/acadgild/install/data/dfs/name/current/fsimage.ckpt_00000000000000000481 (0/0)		
delegation tokens /home/acadgild/install/data/dfs/name/current /fsimage.ckpt_00000000000000000481 (0/0)		
cache pools /home/acadgild/install/data/dfs/name/current /fsimage.ckpt_00000000000000000481 (0/0)		
Safe mode	100%	36 sec
awaiting reported blocks (10/10)		

Hadoop, 2016.

Legacy UI