

ACADGILD

Big Data  
Development  
Assignment 1.1

Ashish Kumar  
ashishashuu0602@gmail.com

---

### **Task 1:**

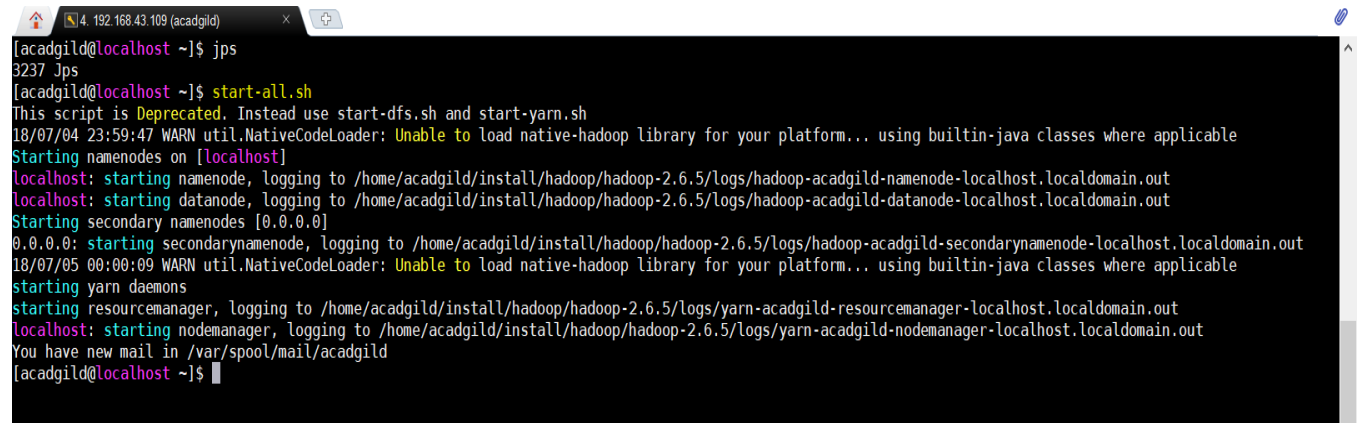
Start Single Node on AcadGild on VM.

**Solution:** start-all.sh

**Description:**

Starts all Hadoop daemons, the namenode, datanode, the resourcemanager, nodemanager, and the secondary namenodes.

**Snapshot:**

A terminal window titled '4. 192.168.43.109 (acadgild)' shows the execution of the 'start-all.sh' script. The prompt is '[acadgild@localhost ~]\$'. The user enters 'jps' and 'start-all.sh'. The script outputs a deprecation warning and then starts various Hadoop daemons: namenode, datanode, secondary namenodes, yarn daemons, resourcemanager, and nodemanager. Each daemon's logging path is displayed. The prompt returns to '[acadgild@localhost ~]\$'.

### **Task 2:**

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

**Solution:** Below is the output of JPS command –

**Description:** JPS command will list out all the running daemons.

```
3363 NameNode
3684 SecondaryNameNode
3828 ResourceManager
3463 DataNode
3930 NodeManager
5021 Jps
```

**Snapshot:**

```
4. 192.168.43.109 (acadgild)
[acadgild@localhost ~]$ jps
3363 NameNode
3684 SecondaryNameNode
3828 ResourceManager
3463 DataNode
3930 NodeManager
5021 Jps
You have new mail in /var/spool/mail/acadgild
```

### **Task 3:**

Run few Unix commands like pwd, ls -ls, etc.

### **Solution:**

- **pwd:**

shows the name of the current working directory.

```
2. 192.168.43.109 (acadgild)
[acadgild@localhost ~]$ pwd
/home/acadgild
You have new mail in /var/spool/mail/acadgild
```

- **ls:**

list files in current working directory except those starting with . and only show the file name.

```
[acadgild@localhost ~]$ ls
Desktop  Downloads  eclipse-workspace  file327.txt  Music  Pictures  Templates  text.txt
Documents  eclipse  employee.java  install  Person.java  Public  test.txt  Videos
```

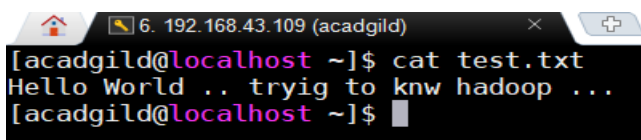
- **ls -l:**

list files in current working directory in long listing format showing permissions, ownership, size and date & timestamp.

```
[acadgild@localhost ~]$ ls -l
total 319576
drwxr-xr-x.  5 acadgild acadgild    4096 Jun 30 12:52 Desktop
drwxr-xr-x.  2 acadgild acadgild    4096 Feb  2 12:52 Documents
drwxr-xr-x.  2 acadgild acadgild    4096 Feb 13 14:24 Downloads
drwxrwxr-x.  3 acadgild acadgild    4096 Dec 29 2017 eclipse
drwxrwxr-x.  3 acadgild acadgild    4096 Jan 16 14:02 eclipse-workspace
-rw-rw-r--.  1 acadgild acadgild 10824 Jun 30 14:02 employee.java
-rw-rw-r--.  1 acadgild acadgild 327155712 Jun 23 17:52 file327.txt
drwxrwxr-x. 13 acadgild acadgild    4096 Feb  9 18:06 install
drwxr-xr-x.  2 acadgild acadgild    4096 Dec 27 2017 Music
-rw-rw-r--.  1 acadgild acadgild 16507 Jun 30 14:45 Person.java
drwxr-xr-x.  2 acadgild acadgild    4096 Dec 27 2017 Pictures
drwxr-xr-x.  2 acadgild acadgild    4096 Dec 27 2017 Public
drwxr-xr-x.  2 acadgild acadgild    4096 Dec 27 2017 Templates
-rw-rw-r--.  1 acadgild acadgild    39 Jun  9 10:48 test.txt
-rw-rw-r--.  1 acadgild acadgild    19 Jun  9 10:20 text.txt
drwxr-xr-x.  2 acadgild acadgild    4096 Dec 27 2017 Videos
```

- **cat:**

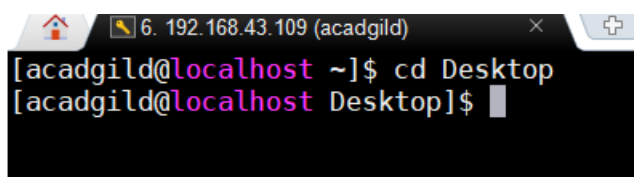
this command lists the content of the small files onto screen/console.



```
6. 192.168.43.109 (acadgild)
[acadgild@localhost ~]$ cat test.txt
Hello World .. tryig to knw hadoop ...
[acadgild@localhost ~]$
```

- **cd:**

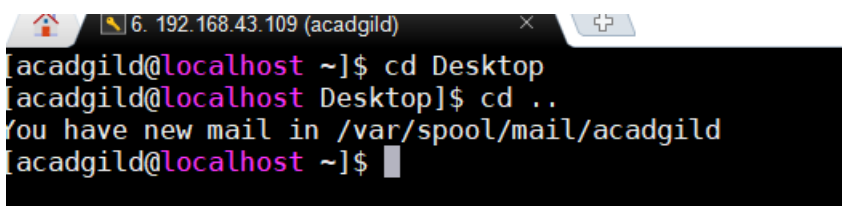
this command changes the current working directory. The '/' indicates path related to root.



```
6. 192.168.43.109 (acadgild)
[acadgild@localhost ~]$ cd Desktop
[acadgild@localhost Desktop]$
```

- **cd ..:**

this command will take you to the parent directory of your current working directory.



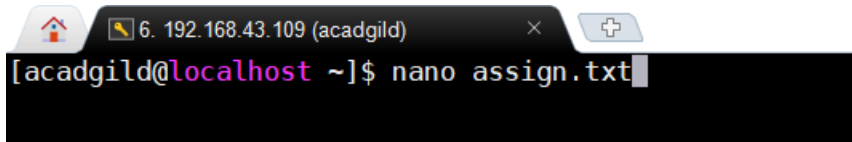
```
6. 192.168.43.109 (acadgild)
[acadgild@localhost ~]$ cd Desktop
[acadgild@localhost Desktop]$ cd ..
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

#### Task 4:

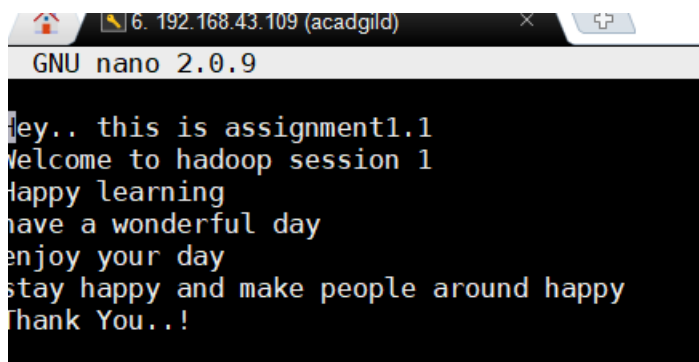
Create a file from the terminal using Nano editor. Add some content and cat it to show content on console.

#### Solution:

`nano <filename>` - this command will create file.



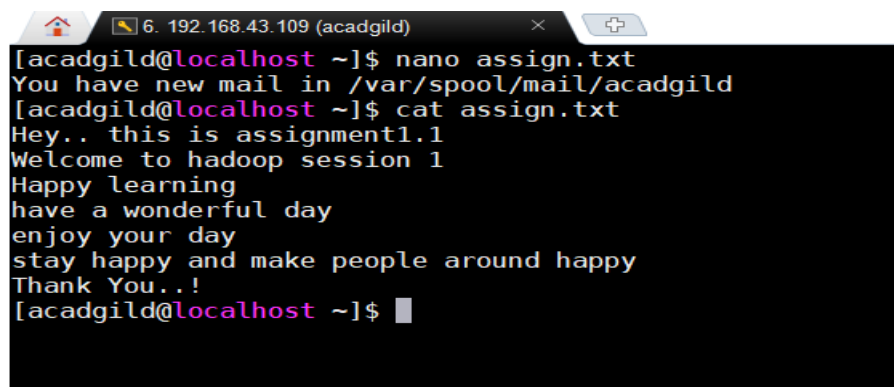
```
[acadgild@localhost ~]$ nano assign.txt
```



```
GNU nano 2.0.9
Hey.. this is assignment1.1
Welcome to hadoop session 1
Happy learning
have a wonderful day
enjoy your day
stay happy and make people around happy
Thank You..!
```

After updating content in the file provide CTRL+X to close editor, then press “Y” to save and then enter to come out of the editor.

To see the content, you can use “cat” command.



```
[acadgild@localhost ~]$ nano assign.txt
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ cat assign.txt
Hey.. this is assignment1.1
Welcome to hadoop session 1
Happy learning
have a wonderful day
enjoy your day
stay happy and make people around happy
Thank You..!
[acadgild@localhost ~]$
```

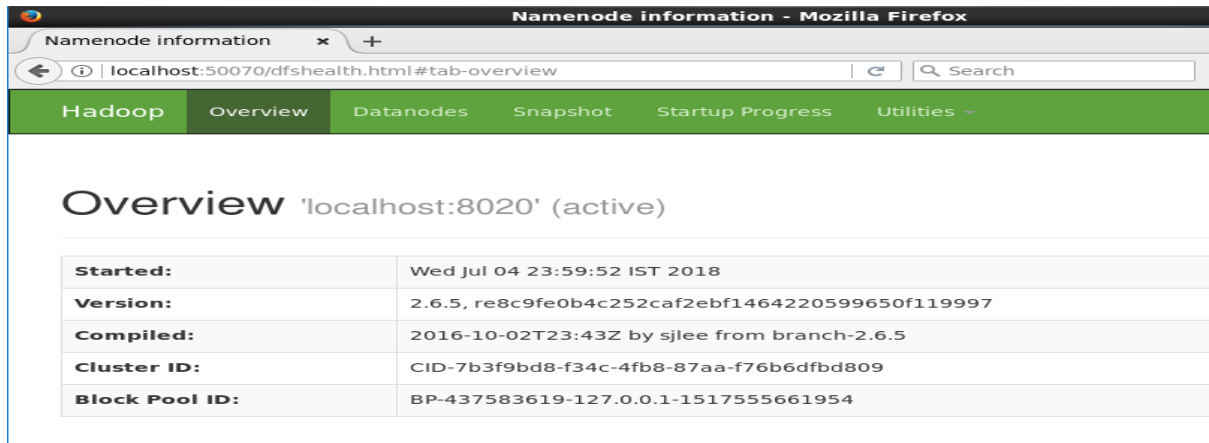
#### Task 5:

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

## Solution:

### Overview tab –

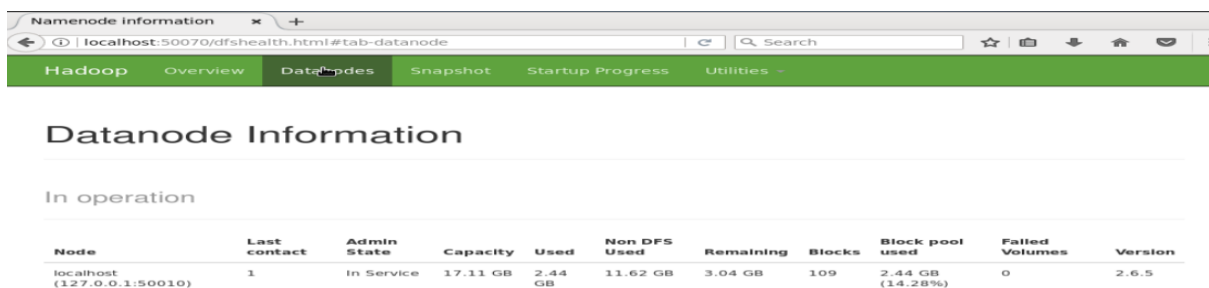
Provides the overview of the cluster like cluster id etc.



<b>Started:</b>	Wed Jul 04 23:59:52 IST 2018
<b>Version:</b>	2.6.5, re8c9fe0b4c252caf2ebf1464220599650f119997
<b>Compiled:</b>	2016-10-02T23:43Z by sjlee from branch-2.6.5
<b>Cluster ID:</b>	CID-7b3f9bd8-f34c-4fb8-87aa-f76b6dfbd809
<b>Block Pool ID:</b>	BP-437583619-127.0.0.1-1517555661954

### Datanode tab –

Provides Datanode Information.



Node	Last contact	Admin State	Capacity	Used	Non DFS Used	Remaining	Blocks	Block pool used	Failed Volumes	Version
localhost (127.0.0.1:50010)	1	In Service	17.11 GB	2.44 GB	11.62 GB	3.04 GB	109	2.44 GB (14.28%)	0	2.6.5

### Snapshot tab –

Snapshot details, this becomes an important feature of Hadoop 2.x as this constitutes the state of the filesystem which brings another level of data protection to HDFS. This is achieved by taking a copy of the filesystem metadata at the point the snapshot is made and making this available to be viewed in the future.

The screenshot shows the 'Snapshot' tab of the Namenode information page. The page title is 'Snapshot Summary'. Below the title, it states 'Snapshottable directories: 0'. There is a table with columns: Path, Snapshot Number, Snapshot Quota, Modification Time, Permission, Owner, and Group. Below this table, it states 'Snapshotted directories: 0'. There is another table with columns: Snapshot ID, Snapshot Directory, and Modification Time. At the bottom, it says 'Hadoop, 2016.' and 'Legacy UI'.

## Startup Progress tab –

This tab provides information about the snapshot components. FSimage and edits information.

The screenshot shows the 'Startup Progress' tab of the Namenode information page. The page title is 'Startup Progress'. Below the title, it states 'Elapsed Time: 36 sec, Percent Complete: 100%'. There is a table with columns: Phase, Completion, and Elapsed Time. The table contains the following data:

Phase	Completion	Elapsed Time
<b>Loading fsimage</b> /home/acadgild/install/data/dfs/name/current /fsimage_00000000000000004894 15.29 KB	100%	0 sec
inodes (0/0)	100%	
delegation tokens (0/0)	100%	
cache pools (0/0)	100%	
<b>Loading edits</b>	100%	0 sec
/home/acadgild/install/data/dfs/name/current /edits_00000000000000004895-00000000000000004895 1 MB (1/1)	100%	
<b>Saving checkpoint</b>	100%	0 sec
inodes /home/acadgild/install/data/dfs/name/current/fsimage.ckpt_00000000000000004895 (0/0)	100%	

## Utilities –

This provides information about each output file. You can check the number of blocks created and also the logs of the job ran. It will provide details with timestamp.

Permission	Owner	Group	Size	Replication	Block Size	Name
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">Hadoopdata</a>
-rw-r--r--	acadgild	supergroup	312 MB	1	128 MB	<a href="#">file327.txt</a>
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">hadoopdata</a>
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">myopt11</a>
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">myou</a>
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">myoutput</a>
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">output</a>
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">output327</a>
drwxr-xr-x	acadgild	supergroup	0 B	0	0 B	<a href="#">sqoopout</a>