

## ASSIGNMENT 2.1

### Task 1: To start Hadoop single node on acadgild VM

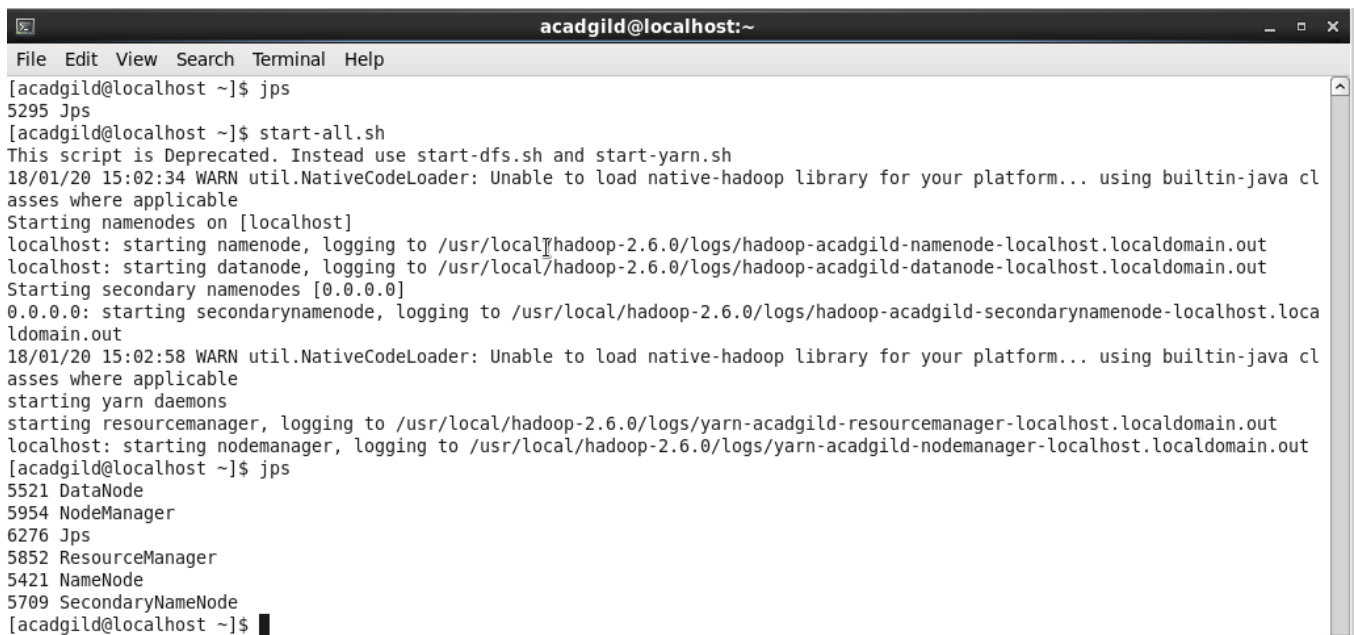
**Command:** start-all.sh

**Explanation:** Above command will start all the daemons. It will start the namenode, datanodes, the jobtracker and tasktrackers in VM.

### Task 2: To check all hadoop daemons are running

**Command:** jps

**Explanation:** It is used to check list of JVMs or daemons running in the system.

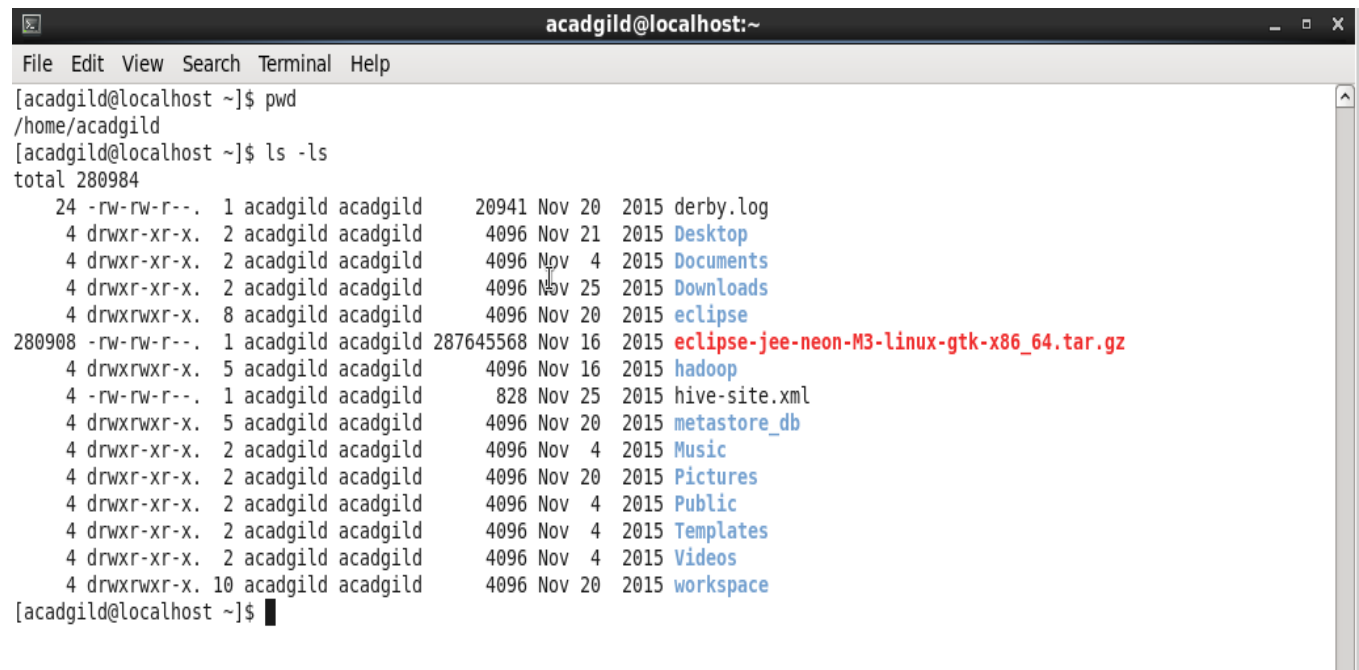
A terminal window titled 'acadgild@localhost:~' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows the execution of 'start-all.sh' and 'jps'. The 'start-all.sh' script outputs deprecation warnings and logs for starting namenode, datanode, secondary namenode, and yarn daemons (resourcemanager, nodemanager). The 'jps' command outputs a list of running JVMs: DataNode (5521), NodeManager (5954), Jps (6276), ResourceManager (5852), NameNode (5421), and SecondaryNameNode (5709).

```
acadgild@localhost:~  
File Edit View Search Terminal Help  
[acadgild@localhost ~]$ jps  
5295 Jps  
[acadgild@localhost ~]$ start-all.sh  
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh  
18/01/20 15:02:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl  
asses where applicable  
Starting namenodes on [localhost]  
localhost: starting namenode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadgild-namenode-localhost.localdomain.out  
localhost: starting datanode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadgild-datanode-localhost.localdomain.out  
Starting secondary namenodes [0.0.0.0]  
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadgild-secondarynamenode-localhost.local  
domain.out  
18/01/20 15:02:58 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl  
asses where applicable  
starting yarn daemons  
starting resourcemanager, logging to /usr/local/hadoop-2.6.0/logs/yarn-acadgild-resourcemanager-localhost.localdomain.out  
localhost: starting nodemanager, logging to /usr/local/hadoop-2.6.0/logs/yarn-acadgild-nodemanager-localhost.localdomain.out  
[acadgild@localhost ~]$ jps  
5521 DataNode  
5954 NodeManager  
6276 Jps  
5852 ResourceManager  
5421 NameNode  
5709 SecondaryNameNode  
[acadgild@localhost ~]$
```

### Task3: Basic commands on AcadGild VM terminal:

**Command:** 1. pwd (It will show the present work directory).

2. ls -ls (It will lists directory contents of files and directories with long format and file size )



```
acadmild@localhost:~  
File Edit View Search Terminal Help  
[acadmild@localhost ~]$ pwd  
/home/acadmild  
[acadmild@localhost ~]$ ls -ls  
total 280984  
  24 -rw-rw-r--.  1 acadmild acadmild    20941 Nov 20  2015 derby.log  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov 21  2015 Desktop  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov  4  2015 Documents  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov 25  2015 Downloads  
   4 drwxr-xr-x.  8 acadmild acadmild     4096 Nov 20  2015 eclipse  
280908 -rw-rw-r--.  1 acadmild acadmild 287645568 Nov 16  2015 eclipse-jee-neon-M3-linux-gtk-x86_64.tar.gz  
   4 drwxrwxr-x.  5 acadmild acadmild     4096 Nov 16  2015 hadoop  
   4 -rw-rw-r--.  1 acadmild acadmild      828 Nov 25  2015 hive-site.xml  
   4 drwxrwxr-x.  5 acadmild acadmild     4096 Nov 20  2015 metastore_db  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov  4  2015 Music  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov 20  2015 Pictures  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov  4  2015 Public  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov  4  2015 Templates  
   4 drwxr-xr-x.  2 acadmild acadmild     4096 Nov  4  2015 Videos  
   4 drwxrwxr-x. 10 acadmild acadmild     4096 Nov 20  2015 workspace  
[acadmild@localhost ~]$
```

### 3. ps aux (display list of all processes)

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~  
[acadgild@localhost ~]$ ps aux  
USER      PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND  
root         1  0.0  0.0  19364  596 ?        Ss   Jan20   0:01 /sbin/init  
root         2  0.0  0.0      0     0 ?        S    Jan20   0:00 [kthreadd]  
root         3  0.0  0.0      0     0 ?        S    Jan20   0:00 [migration/0]  
root         4  0.0  0.0      0     0 ?        S    Jan20   0:01 [ksoftirqd/0]  
root         5  0.0  0.0      0     0 ?        S    Jan20   0:00 [stopper/0]  
root         6  0.0  0.0      0     0 ?        S    Jan20   0:00 [watchdog/0]  
root         7  0.0  0.0      0     0 ?        S    Jan20   0:24 [events/0]  
root         8  0.0  0.0      0     0 ?        S    Jan20   0:00 [events/0]  
root         9  0.0  0.0      0     0 ?        S    Jan20   0:00 [events_long/0]  
root        10  0.0  0.0      0     0 ?        S    Jan20   0:00 [events_power_ef]  
root        11  0.0  0.0      0     0 ?        S    Jan20   0:00 [cgroupl]  
root        12  0.0  0.0      0     0 ?        S    Jan20   0:00 [khelper]  
root        13  0.0  0.0      0     0 ?        S    Jan20   0:00 [netns]  
root        14  0.0  0.0      0     0 ?        S    Jan20   0:00 [async/mgr]  
root        15  0.0  0.0      0     0 ?        S    Jan20   0:00 [pm]  
root        16  0.0  0.0      0     0 ?        S    Jan20   0:00 [sync_supers]  
root        17  0.0  0.0      0     0 ?        S    Jan20   0:00 [bdi-default]  
root        18  0.0  0.0      0     0 ?        S    Jan20   0:00 [kintegrityd/0]  
root        19  0.0  0.0      0     0 ?        S    Jan20   0:07 [kblockd/0]  
root        20  0.0  0.0      0     0 ?        S    Jan20   0:00 [kacpid]  
root        21  0.0  0.0      0     0 ?        S    Jan20   0:00 [kacpi_notify]  
root        22  0.0  0.0      0     0 ?        S    Jan20   0:00 [kacpi_hotplug]  
root        23  0.0  0.0      0     0 ?        S    Jan20   0:00 [ata_aux]  
root        24  0.0  0.0      0     0 ?        S    Jan20   0:11 [ata_sff/0]  
root        25  0.0  0.0      0     0 ?        S    Jan20   0:00 [ksuspend_usbd]  
root        26  0.0  0.0      0     0 ?        S    Jan20   0:00 [khubd]  
root        27  0.0  0.0      0     0 ?        S    Jan20   0:00 [kseriod]  
root        28  0.0  0.0      0     0 ?        S    Jan20   0:00 [md/0]  
root        29  0.0  0.0      0     0 ?        S    Jan20   0:00 [md_misc/0]  
root        30  0.0  0.0      0     0 ?        S    Jan20   0:00 [linkwatch]  
root        32  0.0  0.0      0     0 ?        S    Jan20   0:00 [khungtaskd]  
root        33  0.0  0.0      0     0 ?        S    Jan20   0:04 [kswapd0]  
root        34  0.0  0.0      0     0 ?        SN   Jan20   0:00 [ksmd]  
root        35  0.0  0.0      0     0 ?        SN   Jan20   0:00 [khugepaged]  
root        36  0.0  0.0      0     0 ?        S    Jan20   0:00 [aio/0]
```

### 4.history (it will show all the used commands)

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~  
281 start-all.sh  
282 jps  
283 mkdir /usr/local/hadoop-2.6.0/pid  
284 stop-all.sh  
285 start-all.sh  
286 vi /usr/local/hadoop-2.6.0/etc/hadoop/hadoop-env.sh  
287 stop-all.sh  
288 start-all.sh  
289 jps  
290 hadoop fs -ls /  
291 echo $HADOOP_PID_DIR  
292 vi /usr/local/hadoop-2.6.0/etc/hadoop/hadoop-env.sh  
293 stop-all.sh  
294 jps  
295 cat /usr/local/hadoop-2.6.0/etc/hadoop/core-site.xml  
296 vi /usr/local/hadoop-2.6.0/etc/hadoop/core-site.xml  
297 cat /usr/local/hadoop-2.6.0/etc/hadoop/hdfs-site.xml  
298 wq!  
299 jps  
300 kill -9 3336 3434 3336 3622  
301 jps  
302 start-all.sh  
303 jps  
304 hadoop fs -ls /  
305 cd hadoop/pid/  
306 ls  
307 sudo vi i  
308 ls -l /etc/init.d/  
309 ls -l /etc/init.d/ha*  
310 ls -l /etc/init.d/had*  
311 sudo rm -r /etc/init.d/hadoop-service  
312 cd  
313 sudo cp hadoop-service hadoop-service.sh  
314 ls -l  
315 ls -l /etc/init.d/  
316 ls -l /usr/local/hadoop-2.6.0/  
317 vi hadoop-service
```

5.touch (command use to create blank file)

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~  
[acadgild@localhost ~]$ touch sample2.txt  
[acadgild@localhost ~]$ ls -lh  
total 275M  
-rw-rw-r--. 1 acadgild acadgild 21K Nov 20 2015 derby.log  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 21 2015 Desktop  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 4 2015 Documents  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 25 2015 Downloads  
drwxrwxr-x. 8 acadgild acadgild 4.0K Nov 20 2015 eclipse  
-rw-rw-r--. 1 acadgild acadgild 275M Nov 16 2015 eclipse-jee-neon-M3-linux-gtk-x86_64.tar.gz  
-rw-rw-r--. 1 acadgild acadgild 41 Jan 20 18:39 file1.txt  
-rw-rw-r--. 1 acadgild acadgild 41 Jan 20 18:51 file2.txt  
drwxrwxr-x. 2 acadgild acadgild 4.0K Jan 21 03:21 folder1  
drwxrwxr-x. 5 acadgild acadgild 4.0K Nov 16 2015 hadoop  
-rw-rw-r--. 1 acadgild acadgild 828 Nov 25 2015 hive-site.xml  
-rw-rw-r--. 1 acadgild acadgild 135 Jan 21 00:09 max-temp.txt  
drwxrwxr-x. 5 acadgild acadgild 4.0K Nov 20 2015 metastore_db  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 4 2015 Music  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 20 2015 Pictures  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 4 2015 Public  
-rw-rw-r--. 1 acadgild acadgild 557 Jan 21 03:28 sample1.txt  
-rw-rw-r--. 1 acadgild acadgild 0 Jan 21 03:49 sample2.txt  
-rw-rw-r--. 1 acadgild acadgild 573 Jan 20 23:49 sample.txt  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 4 2015 Templates  
drwxr-xr-x. 2 acadgild acadgild 4.0K Nov 4 2015 Videos  
drwxrwxr-x. 10 acadgild acadgild 4.0K Nov 20 2015 workspace  
[acadgild@localhost ~]$
```

**Task4:To create a test document using nano command**

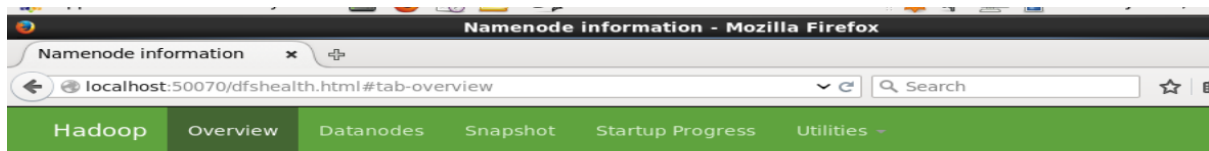
**Command:** nano sample1.txt

Cat sample1.txt

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~  
[acadgild@localhost ~]$ nano sample1.txt  
[acadgild@localhost ~]$ cat sample1.txt  
The Apache Hadoop software library is a framework that allows for the distributed processing of large data sets across  
clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of  
machines, each offering local computation and storage. Rather than rely on hardware to deliver high-availability,  
the library itself is designed to detect and handle failures at the application layer, so delivering a highly-available  
service on top of a cluster of computers, each of which may be prone to failures.  
[acadgild@localhost ~]$
```

**Task5: To open the hdfs web page localhost:50070 and to check details of hdfs.**

**Output-**



## Overview 'localhost:9000' (active)

Started:	Sat Jan 20 15:02:40 IST 2018
Version:	2.6.0, re3496499ecb8d220fba99dc5ed4c99c8f9e33bb1
Compiled:	2014-11-13T21:10Z by jenkins from (detached from e349649)
Cluster ID:	CID-71615a0b-f560-40e7-94e0-1057e6718ac8
Block Pool ID:	BP-695350366-127.0.0.1-1445620353823

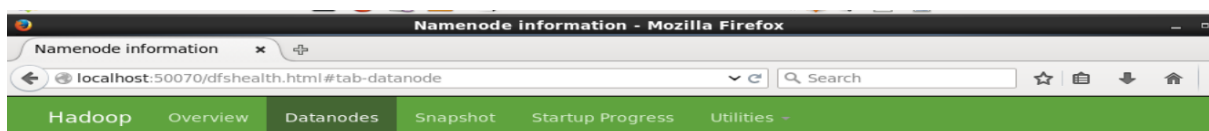
## Summary

Security is off.

Safemode is off.

228 files and directories, 164 blocks = 392 total filesystem object(s).

Heap Memory used 44.13 MB of 57.42 MB Heap Memory. Max Heap Memory is 966.69 MB.



## 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)	1	In Service	17.11 GB	103.73 MB	7.7 GB	9.31 GB	164	103.73 MB (0.59%)	0	2.6.0

### Decommissioning

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