Assignment Submission- Session 1
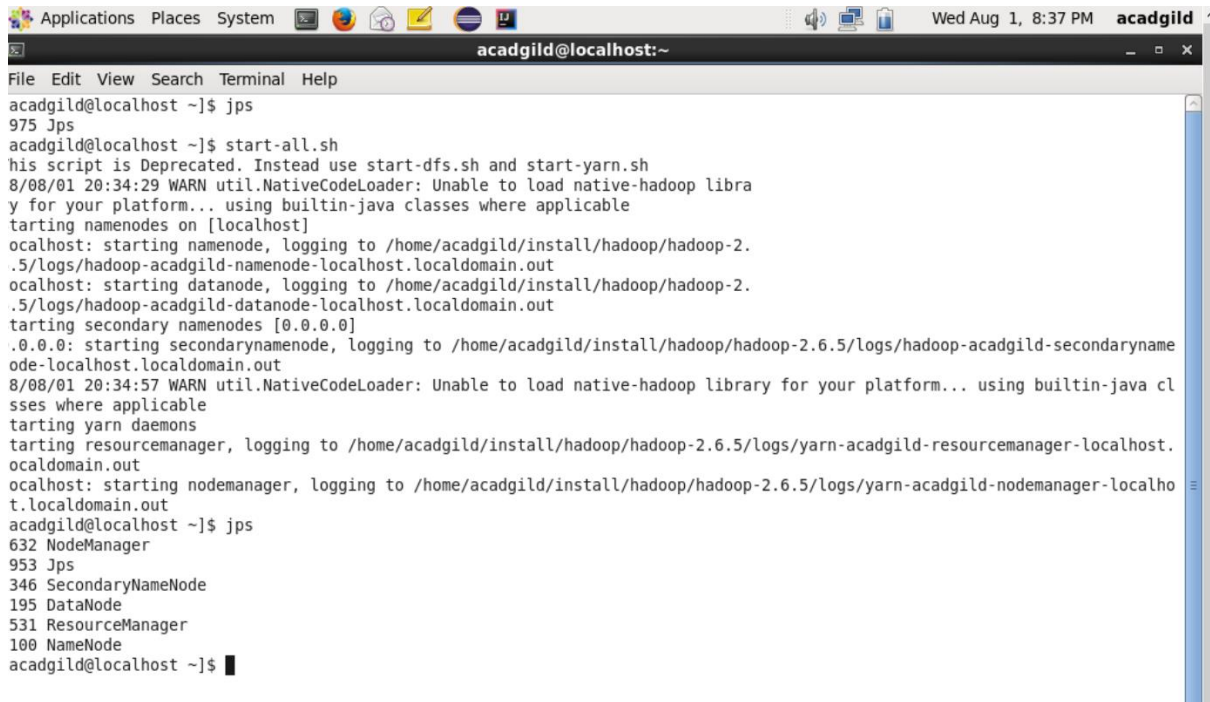
1.) **Start Hadoop Single Node.**
2.) **Run jps command to see if all Hadoop daemons are running.**

JPS gives current "Java Process Status"

After running command "start-all.sh". The Hadoop daemons got started. The above screenshot shows processes started along with their process ids.

"Resource manager" and "NameNode" are the master daemons, where as "Secondary NameNode" is kind of assistant manager to "NameNode" which keeps track of "fsImage/namespace image" file of the "NameNode" and updates the same with "edits" and provide it back to the "NameNode". It keeps latest copy of "namespace file".

"Node Manager" and "Data Node" are the slave daemons that run on all the machines in the cluster, where actual data storage and processing happens on the HDFS cluster.

3.) Run a few unix commands like pwd, ls etc.

```
                                        acadgild@localhost:~                                    _ □ X
File  Edit  View  Search  Terminal  Help
[acadgild@localhost ~]$ pwd
/home/acadgild
[acadgild@localhost ~]$ ls -l
total 72
-rw-rw-r--.  1 acadgild acadgild   75 Jul 26 11:32 aks_test
-rw-rw-r--.  1 acadgild acadgild   75 Jul 26 11:31 aks_test~
-rw-r--r--.  1 acadgild acadgild   75 Jul 26 11:48 aks_test.txt
drwxr-xr-x.  3 acadgild acadgild 4096 May 15 17:57 Desktop
drwxr-xr-x.  2 acadgild acadgild 4096 May 14 14:05 Documents
drwxr-xr-x.  4 acadgild acadgild 4096 May 16 17:15 Downloads
drwxrwxr-x.  3 acadgild acadgild 4096 May 15 16:13 eclipse
drwxrwxr-x.  3 acadgild acadgild 4096 May 15 16:16 eclipse-workspace
drwxrwxr-x.  3 acadgild acadgild 4096 May 16 17:24 IdeaProjects
drwxrwxr-x. 13 acadgild acadgild 4096 May 16 13:33 install
drwxrwxr-x.  8 acadgild acadgild 4096 May 16 17:16 IntelliJ
drwxrwxr-x.  5 acadgild acadgild 4096 May 16 13:41 metastore_db
drwxr-xr-x.  2 acadgild acadgild 4096 May 14 14:05 Music
drwxr-xr-x.  2 acadgild acadgild 4096 May 14 14:05 Pictures
drwxr-xr-x.  2 acadgild acadgild 4096 May 14 14:05 Public
drwxr-xr-x.  2 acadgild acadgild 4096 May 14 14:05 Templates
-rw-rw-r--.  1 acadgild acadgild 2435 May 22 13:36 tweets
drwxr-xr-x.  2 acadgild acadgild 4096 May 14 14:05 Videos
[acadgild@localhost ~]$ hadoop fs -ls /
18/08/01 20:54:18 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 3 items
-rw-r--r--   1 acadgild supergroup         75 2018-07-26 11:38 /aks_test.txt
drwxr-xr-x   - acadgild supergroup          0 2018-05-16 13:49 /hbase
drwx-wx-wx   - acadgild supergroup          0 2018-05-16 13:42 /tmp
[acadgild@localhost ~]$ █
```

→"pwd" shows present working directory.

→ "ls -l" lists all the files/sub directories present in the pwd along with details. It doesn't show hidden files/directories. (ls -la can be used to see all)
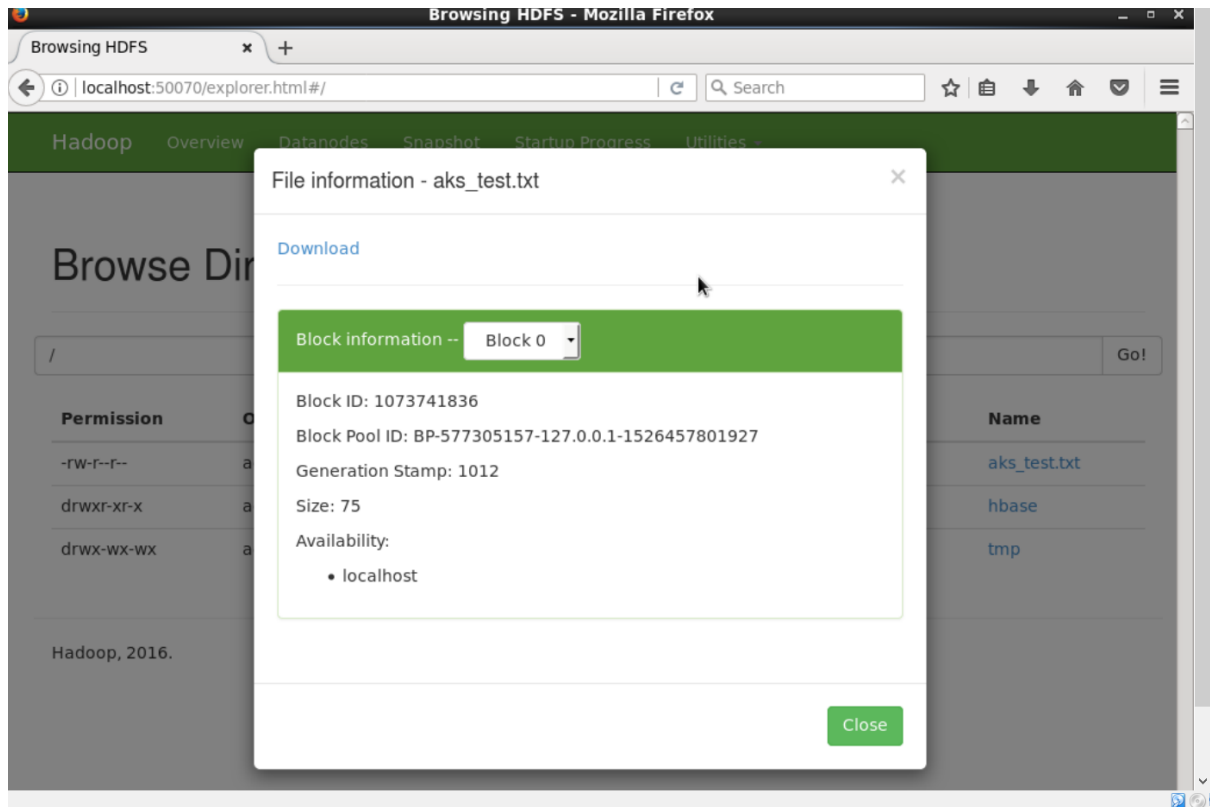
- -rw*** signifies it's a file, not a directory, drw-*** signifies it's a directory. Also signifies access user has on files (read/write).

→ "hadoop fs -ls" shows all the files/directories present in the hdfs.

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

```
File  Edit  View  Search  Terminal  Help
[acadgild@localhost ~]$ pwd
/home/acadgild
[acadgild@localhost ~]$ mkdir akshat_dir
[acadgild@localhost ~]$ cd akshat_dir
[acadgild@localhost akshat_dir]$ pwd
/home/acadgild/akshat_dir
[acadgild@localhost akshat_dir]$ nano assignment_1.txt
[acadgild@localhost akshat_dir]$ pwd
/home/acadgild/akshat_dir
[acadgild@localhost akshat_dir]$ ls -la
total 12
drwxrwxr-x.  2 acadgild acadgild 4096 Aug  1 21:17 .
drwx------. 49 acadgild acadgild 4096 Aug  1 21:09 ..
-rw-rw-r--.  1 acadgild acadgild   76 Aug  1 21:17 assignment_1.txt
[acadgild@localhost akshat_dir]$ cat assignment_1.txt
This is my first created file using a nano editor. Creted for assignment 1.
[acadgild@localhost akshat_dir]$ █
```

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

Block information of the file moved to hdfs.