**Task 1:**

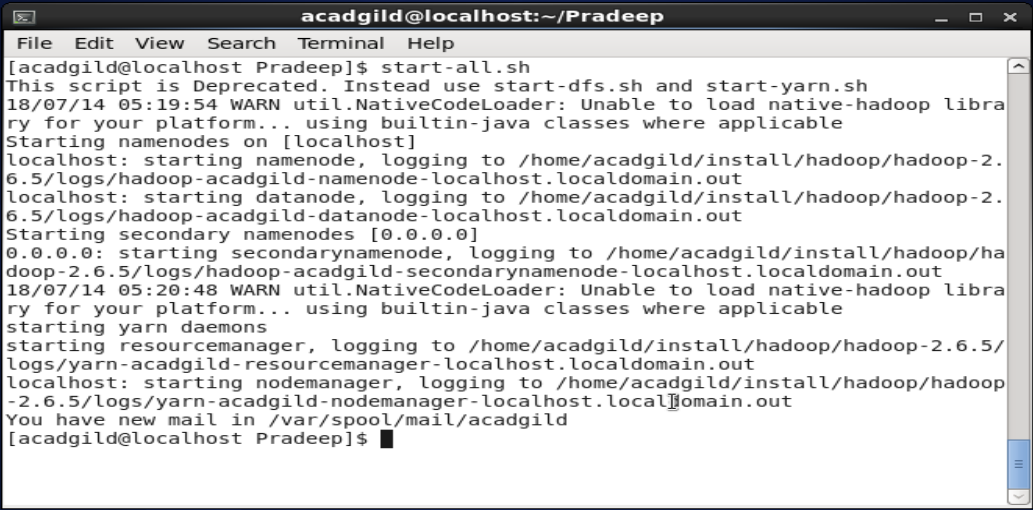
1. Start Hadoop single node on AcadGild VM (The command is start-all.sh)

**OUTPUT:**

Start-all.sh will start the demons required for running HDFS and MapReduce programs/applications.

This command will start the following demons.

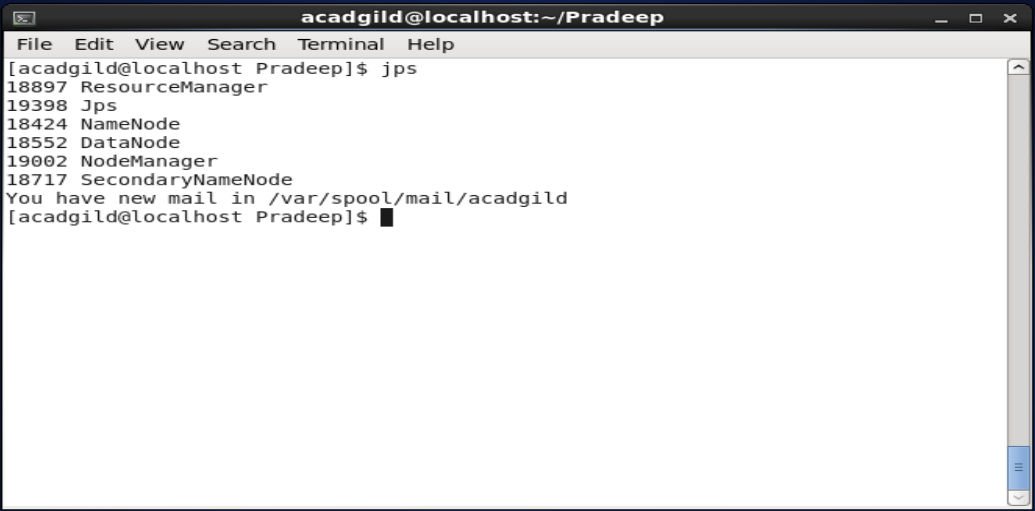
* NameNode
* SecondaryNameNode
* DataNode
* ResourceManager
* DataManager



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

**OUTPUT:**

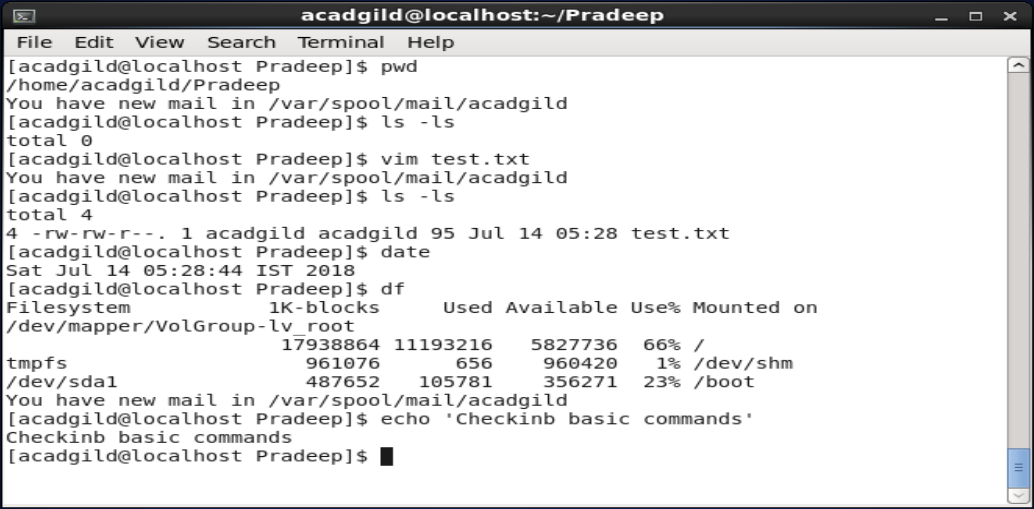
jps is the command to check what all java process running. Before we start any hadoop operations we need to make sure all the HDFS and MapReduce demons/processes running.



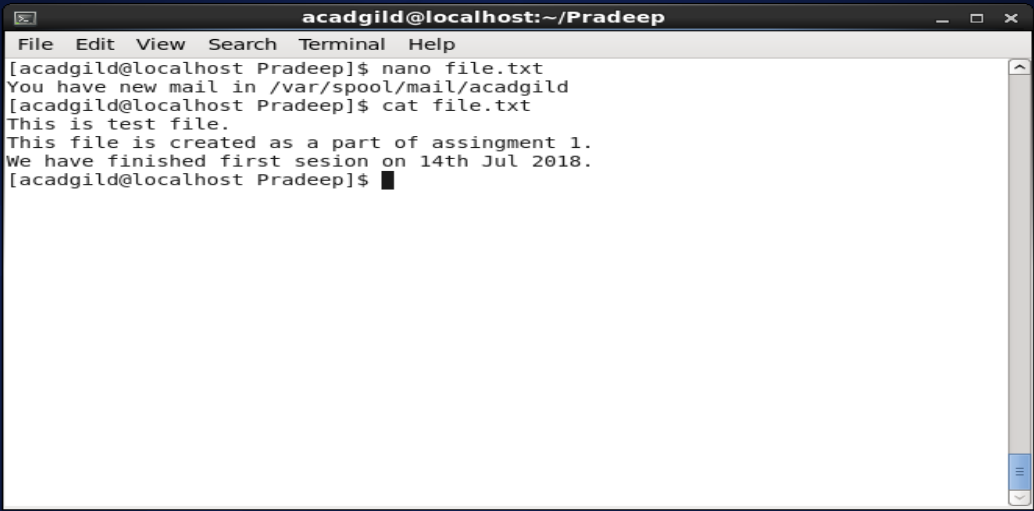
3. Run a couple of Unix commands like pwd, ls -ls etc.

**OUTPUT:**

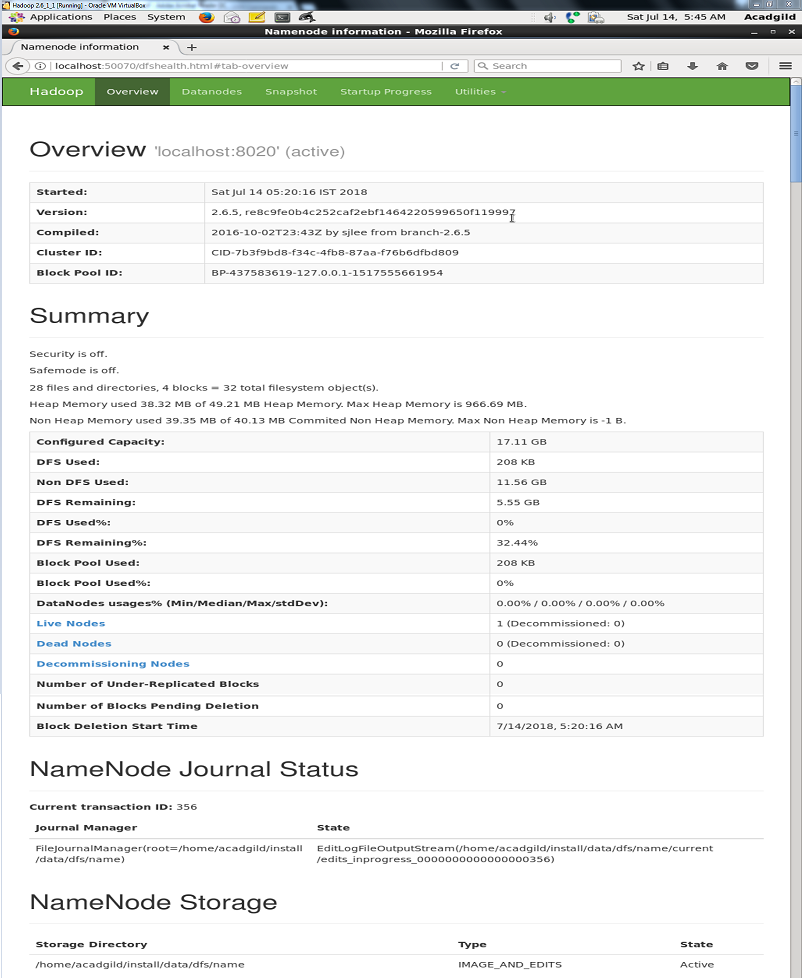
1. pwd –> prints the present working directory where you have fired this command
2. ls –ls –> list the contents of the folder with **l**ong format and **s**ize allotted for each item
3. vim test.txt -> opens the editor for creating a file named test.txt and you can add the contents and save and close



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



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



**Task 2:**

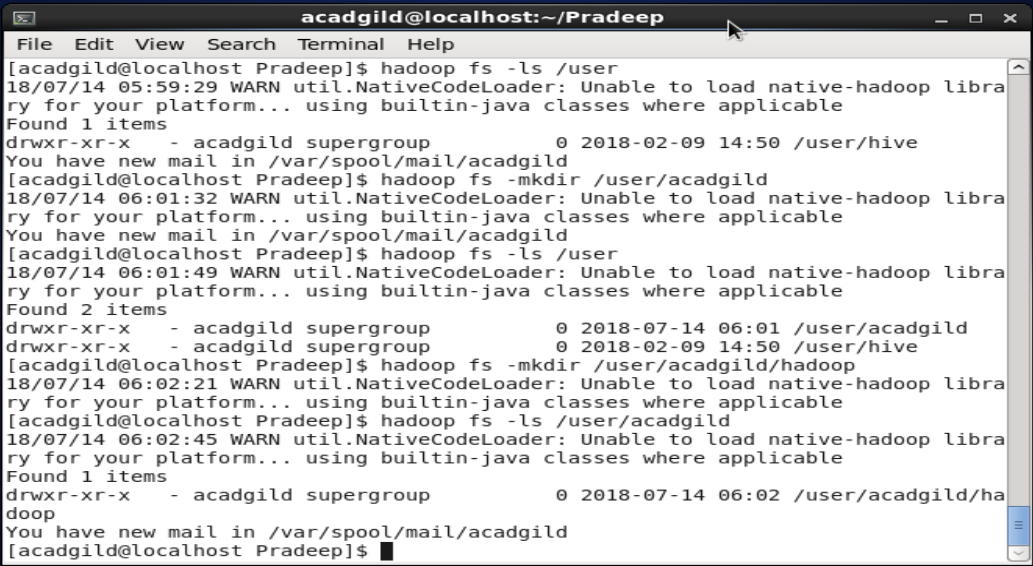
1. Check whether /user/acadgild directory exists or not in the HDFS.

If it doesn't exist, then create this.

Create a directory /user/acadgild/hadoop.

**OUTPUT:**

1. First checking the available folders under /user with the following command:
   1. Hadoop fs –ls /user
2. There was no acadgild folder under /user so created this folder with the following command:
   1. Hadoop fs –mkdir /user/acadgild
3. Then created the hadoop folder with in acadgild folder with the following command:
   1. Hadoop fs –mkdir /user/acadgild/hadoop
4. Then verified the whether operations are successful or not by the following command:
   1. Hadoop fs –ls /user/acadgild/hadoop



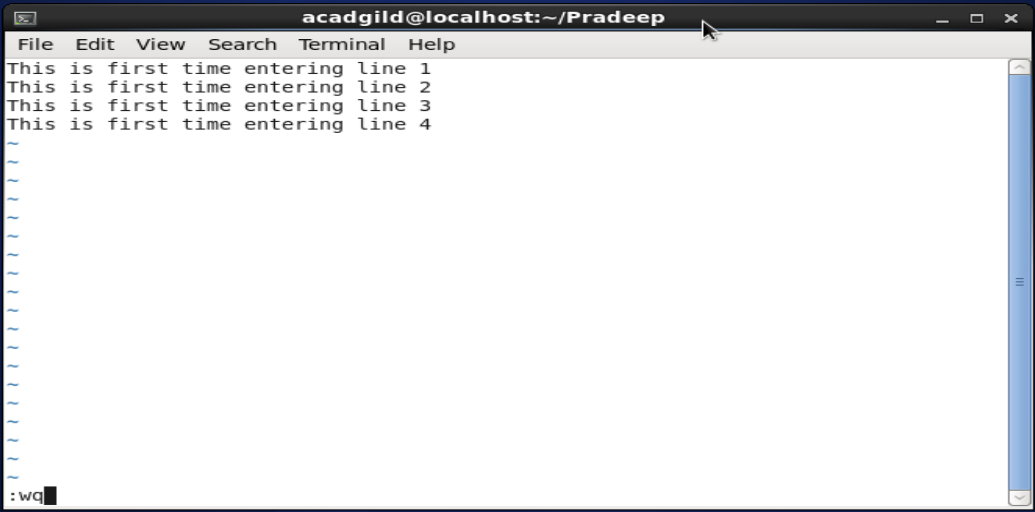
2. Create a file in HDFS under directory /user/acadgild/hadoop, with name word-count.txt.

Whatever we type on screen should get appended to the file. Try to type (on screen) few lines

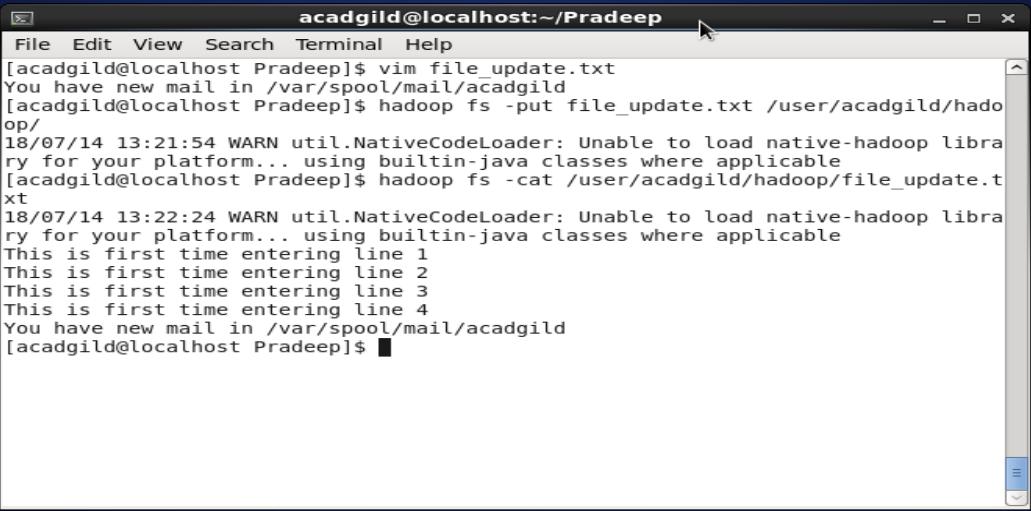
from any online article or textbook.

**OUPUT:**

1. I have created the file\_update.txt file in local file system with the following contents:



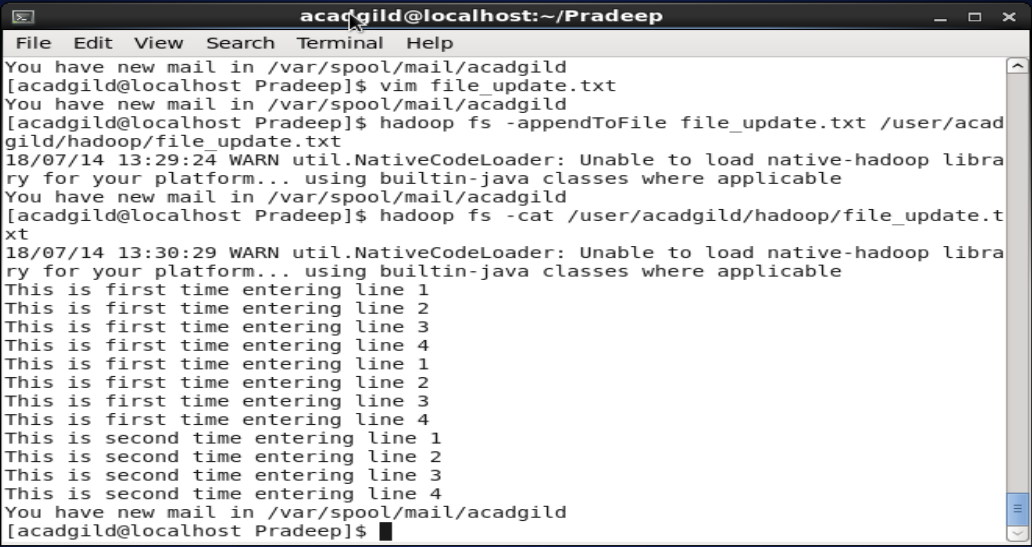
1. Then I have put this file in hadoop file structure with the following command:
   1. Hadoop fs –put file\_update.txt /user/acadgild/hadoop/
2. Then I have displayed the content of the file\_update.txt of hadoop FS with the following command:
   1. Hadoop fs –cat /user/acadgild/hadoop/file\_update.txt

****

1. Now I have added some more content in the file\_update.txt file present in the local FS with the following contents:

****

1. Then I have used the appendFileTo command to append this content to already present file\_update.txt file in hadoop FS by the following command:
   1. Hadoop fs –appendToFile file\_update.txt /user/acadgild/hadoop/file\_update.txt
2. Then I verified the file\_update.txt that is present in hadoop FS whether the new contents are appended to the file or not by hadoop fs –cat command as shown in screen shot:

****

3. Create a file max-temp.txt in local FS. Put some 10-15 records of date and temperature

example: dd-mm-yyyy,temperature

Example:

10-01-1990,10

10-02-1991,20

Move this file to HDFS at /user/acadgild/hadoop.

**OUTPUT:**

1. I have used vim command to create the max\_temprature.txt file with some content in local file system (/home/acadgild/Pradeep/).
2. I have used the following command to copy this file to Hadoop File System (/user/acadgild/hadoop/).
   1. Hadoop fs –put /home/acadgild/Pradeep/max\_temperature /user/acadgild/hadoop/
   2. Verified whether file is copied or not with following command:
      1. Hadoop fs –ls /user/acadgild/hadoop
3. Screen shot of these operations attached below:



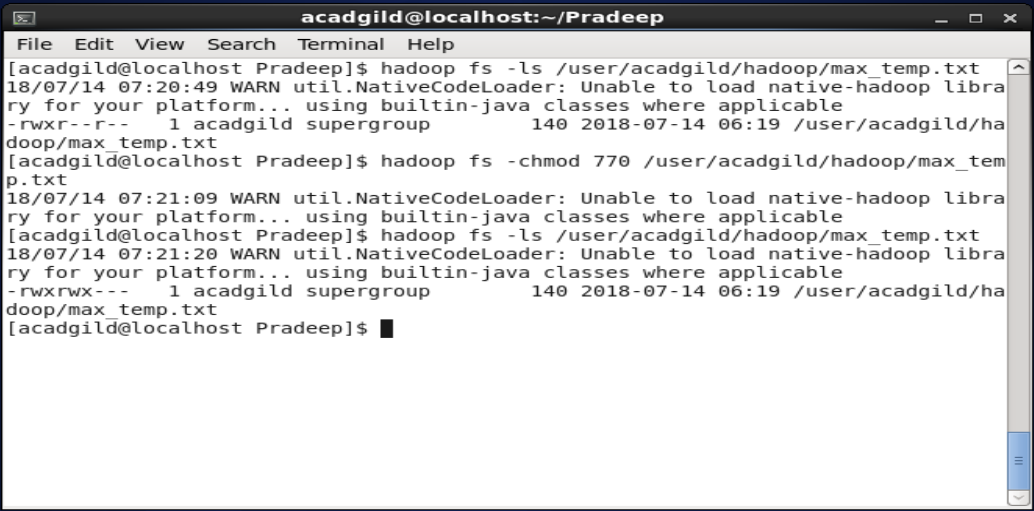
4. Change the permission of the file /user/acadgild/hadoop/max-temp.txt, such that only the

owner and the group members have full control over the file.

Others do not have any control over it.

OUTPUT:

1. First checked the current permissions of the max\_temp.txt file with the following command:
   1. Hadoop fs –ls /user/acadgild/hadoop/max\_temp.txt
2. Changed the user and group access with full control by the following command:
   1. Hadoop fs –chmod 770 /user/acadgild/hadoop/max\_temp.txt
3. Again verified the permissions of the max\_temp.txt file with the following command to verify changes:
   1. Hadoop fs –ls /user/acadgild/hadoop/max\_temp.txt
4. Screen shot of the above operations attached below:



**Task 3:**

1. Execute **WordMedian** , **WordMean** , **WordStandardDeviation** programs using

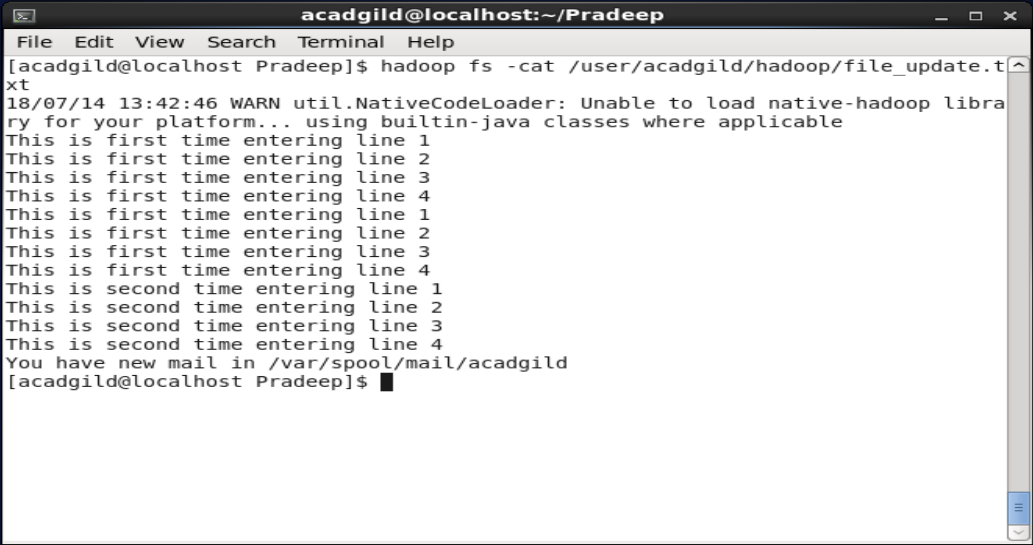
hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.

Refer path below.

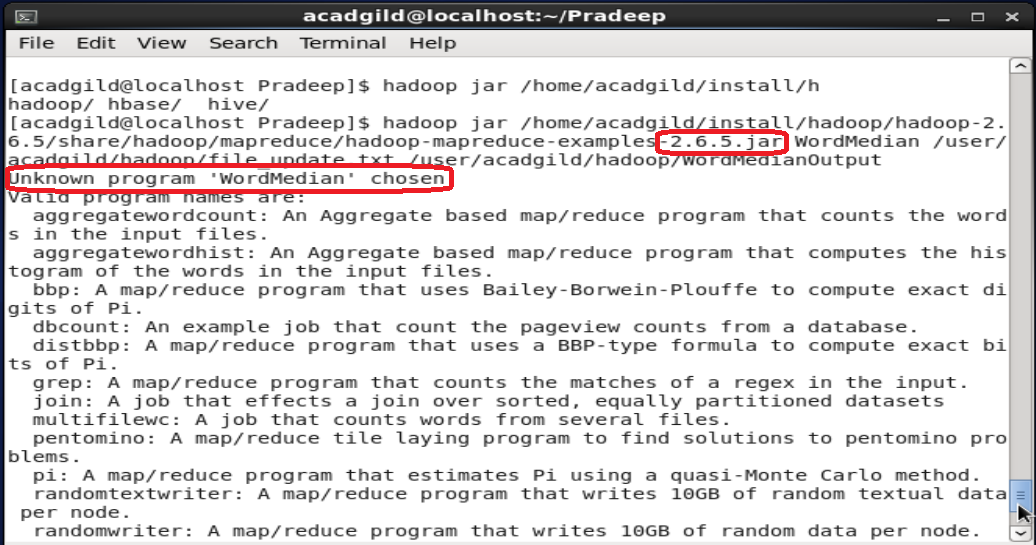
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce

**OUTPUT:**

1. I will use the following hadoop file system file file\_update.txt for execute the jar file:



1. Then I have executed the following command:
   1. Hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/ hadoop-mapreduce-examples-2.6.5.jar WordMedian /user/acadgild/hadoop/file\_update.txt /user/acadgild/hadoop/WordMedianOutput



1. In VM I don’t see hadoop-mapreduce-example-2.9.0.jar file so couldn’t able to execute the above command successfully.
2. But I understood the problem statement and commands to achieve the given classes like WordMedian, WordMean, WordStandardDeviation results and store the output in respective output folders like WordMedianOutput, WordMeanOutput, WordStandardDeviationOutput