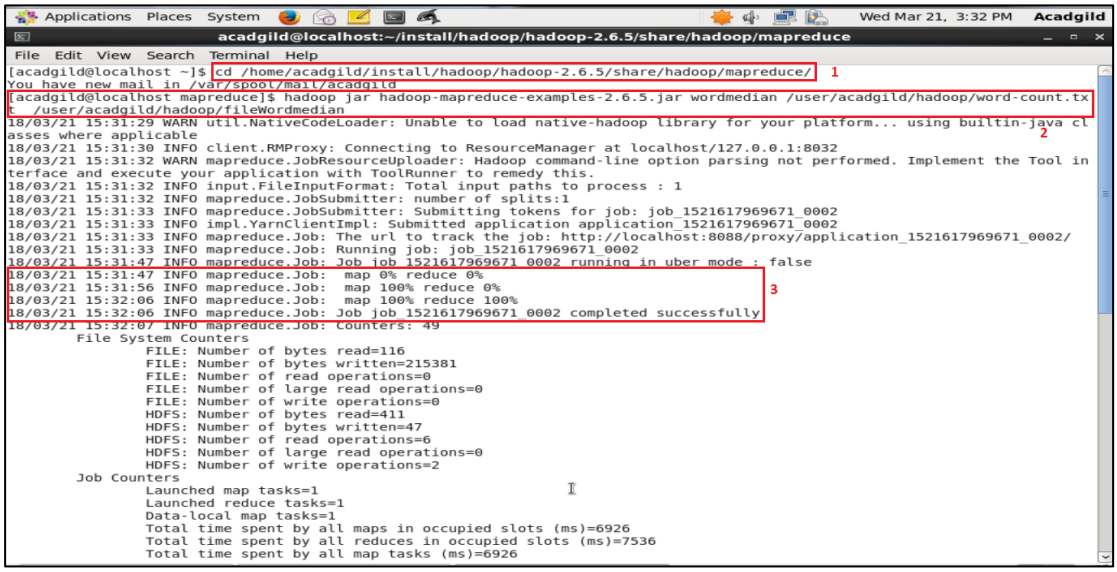
## Session 3 - YARN

## Assignment 1

**Task1 :** Execute **WordMedian** , **WordMean** , **WordStandardDeviation** programs using hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.

WordMedian :




With reference to the screenshot above,

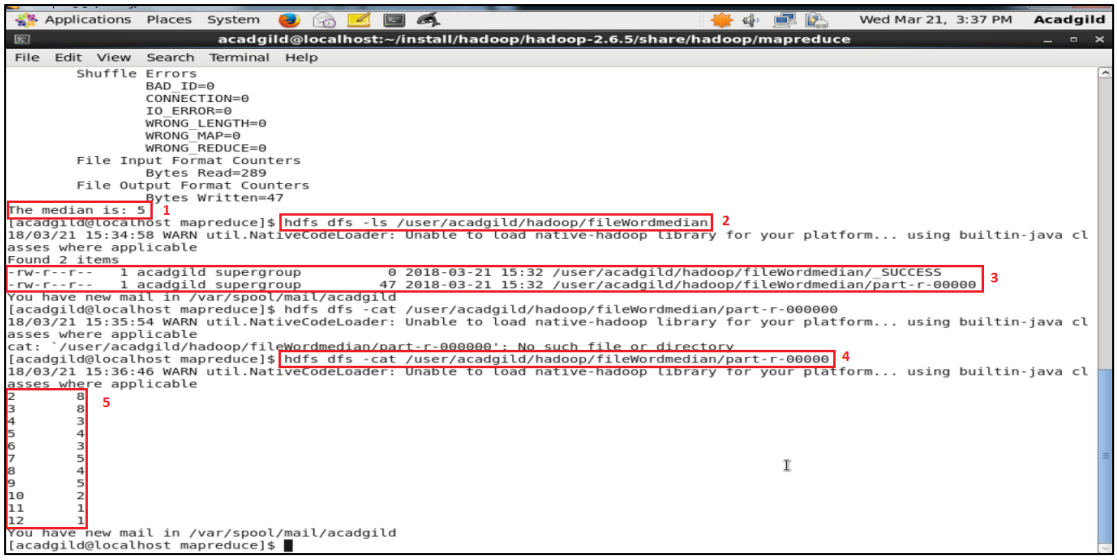
1 : Move to the folder where the jar file is located using the following command :

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

2 : Execute the WordMedian class from the jar file by passing 2 parameter to it : the source location of the file and a location for the target files to be stored. So the command will be :

*hadoop jar hadoop-mareduce-examples-2.6.5.jar wordmedian /user/acadgild/hadoop/word-count.txt /user/acadgild/hadoop/fileWordmedian*

3 : As the jar is executed a map reduce job is initiated. The mapper is executed completely first and then the reducer job is picked up. Once both are executed, the job is said to be completed successfully.




With reference to the screenshot above,

1 : Once the job is completed, the output is printed as follows :

The median is: 5

2 : When navigated to the location given in the hadoop jar command, there are 2 files that are created. The command used is

*hdfs dfs -ls /user/acadgild/hadoop/fileWordmedian*

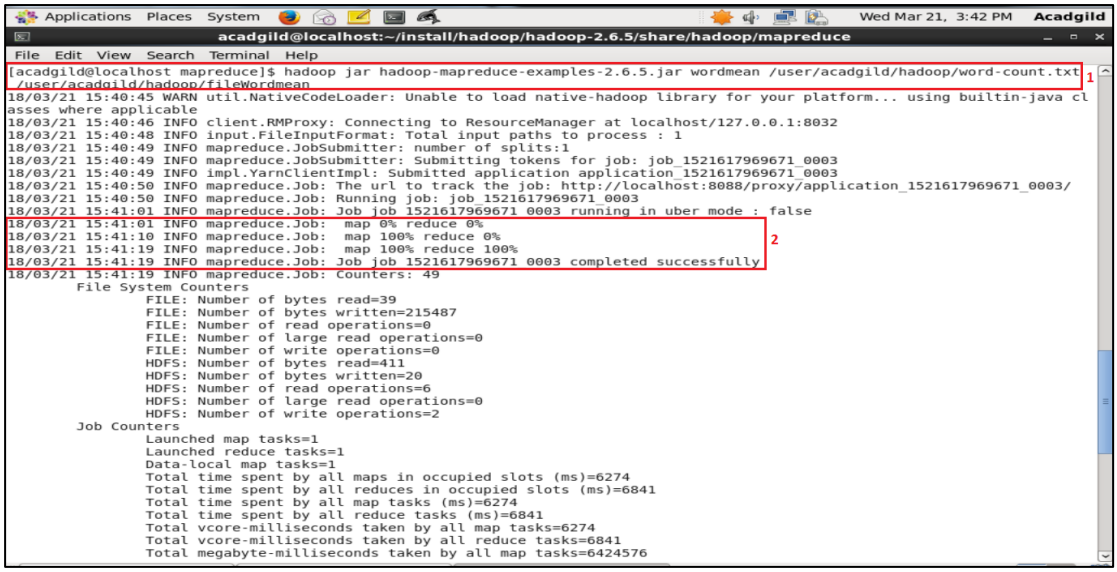
3 : _SUCCESS : indicating that the job is completed successfully.

part-r-00000 : the part file (a partition) that contains the output of the job.

4 : To view the contents of the part file, a cat command is used.

*hdfs dfs -cat /user/acadgild/hadoop/fileWordmedian/part-r-00000*

WordMean :

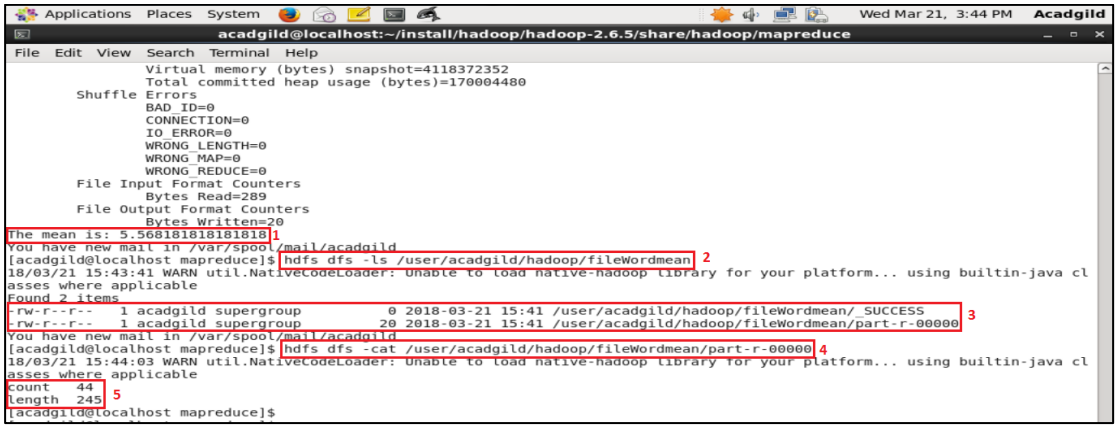



With reference to the screenshot above,

1 : Execute the WordMean class from the jar file by passing 2 parameter to it : the source location of the file and a location for the target files to be stored. So the command will be :

*hadoop jar hadoop-mareduce-examples-2.6.5.jar wordmean /user/acadgild/hadoop/word-count.txt /user/acadgild/hadoop/fileWordmean*

2 : As the jar is executed a map reduce job is initiated. The mapper is executed completely first and then the reducer job is picked up. Once both are executed, the job is said to be completed successfully.




With reference to the screenshot above,

1 : Once the job is completed, the output is printed as follows :

The mean is: 5.5681818181818181

2 : When navigated to the location given in the hadoop jar command, there are 2 files that are created. The command used is

*hdfs dfs -ls /user/acadgild/hadoop/fileWordmean*

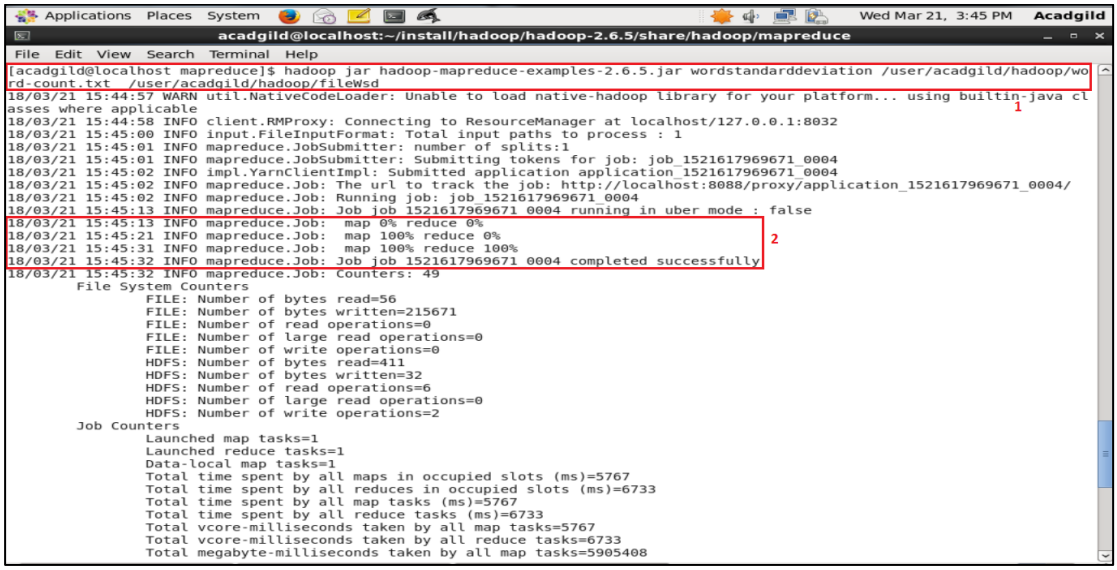
3 : _SUCCESS : indicating that the job is completed successfully.

part-r-00000 : the part file (a partition) that contains the output of the job.

4 : To view the contents of the part file, a cat command is used.

*hdfs dfs -cat /user/acadgild/hadoop/fileWordmean/part-r-00000*

WordStandardDeviation :

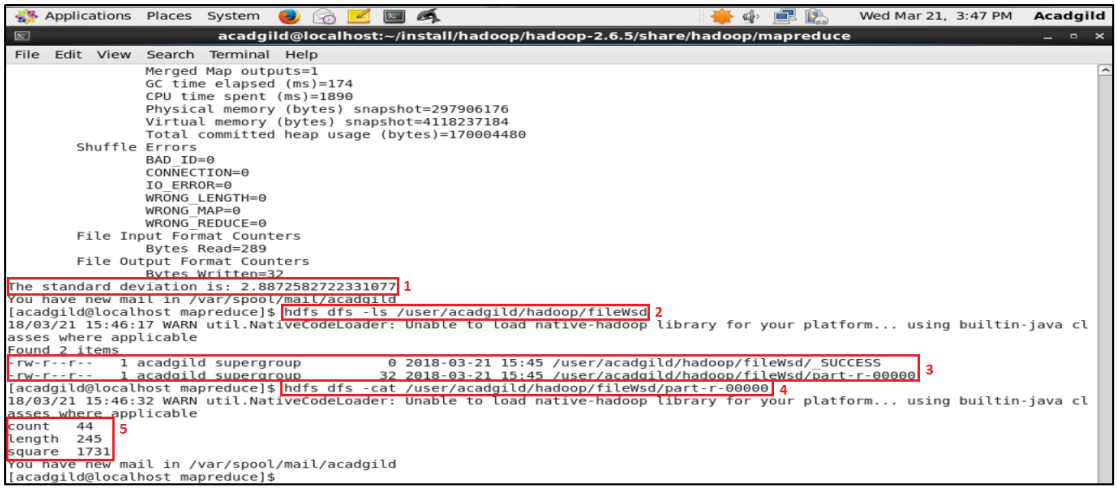



With reference to the screenshot above,

1 : Execute the WordStandardDeviation class from the jar file by passing 2 parameter to it : the source location of the file and a location for the target files to be stored. So the command will be :

*hadoop jar hadoop-mareduce-examples-2.6.5.jar wordstandarddeviation /user/acadgild/hadoop/word-count.txt /user/acadgild/hadoop/fileWsd*

2 : As the jar is executed a map reduce job is initiated. The mapper is executed completely first and then the reducer job is picked up. Once both are executed, the job is said to be completed successfully.




With reference to the screenshot above,

1 : Once the job is completed, the output is printed as follows :

The standard deviation is: 2.8872582722331077

2 : When navigated to the location given in the hadoop jar command, there are 2 files that are created. The command used is

*hdfs dfs -ls /user/acadgild/hadoop/fileWsd*

3 : _SUCCESS : indicating that the job is completed successfully.

part-r-00000 : the part file (a partition) that contains the output of the job.

4 : To view the contents of the part file, a cat command is used.

*hdfs dfs -cat /user/acadgild/hadoop/fileWsd/part-r-00000*