

# Session 3: YARN

## Assignment 1

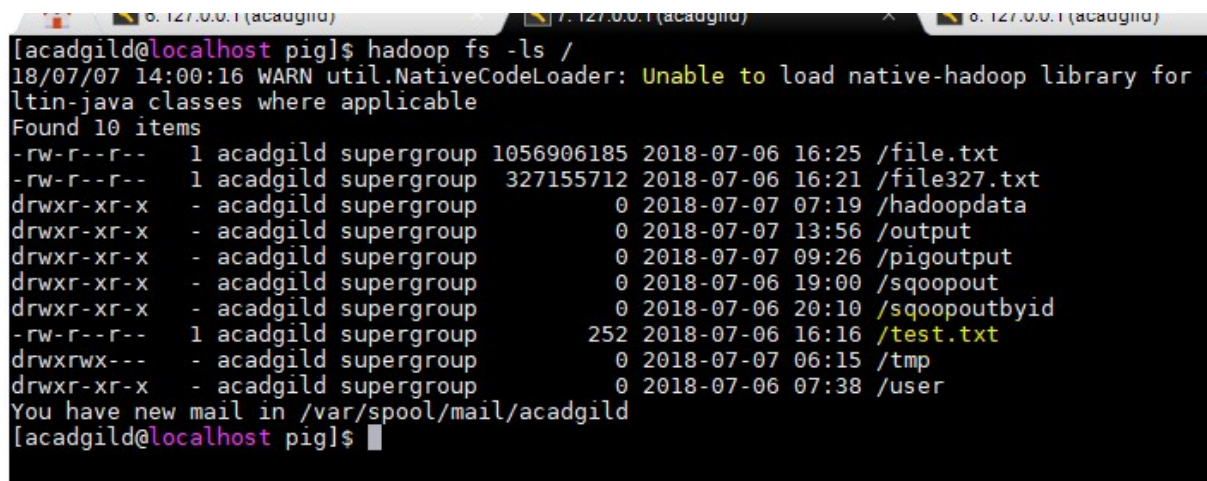
### Task 1:

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

Step:1 – Check test.txt file is present in HDFS.

A terminal window screenshot showing the command 'hadoop fs -ls /' being executed. The output lists 10 items in HDFS, including /file.txt, /file327.txt, /hadoopdata, /output, /pigoutput, /sqoopout, /sqoopoutbyid, /test.txt (highlighted in yellow), /tmp, and /user. A warning message about the native-hadoop library is also visible at the top of the output.

```
[acadgild@localhost pig]$ hadoop fs -ls /
18/07/07 14:00:16 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
ltin-java classes where applicable
Found 10 items
-rw-r--r--  1 acadgild supergroup 1056906185 2018-07-06 16:25 /file.txt
-rw-r--r--  1 acadgild supergroup  327155712 2018-07-06 16:21 /file327.txt
drwxr-xr-x  - acadgild supergroup           0 2018-07-07 07:19 /hadoopdata
drwxr-xr-x  - acadgild supergroup           0 2018-07-07 13:56 /output
drwxr-xr-x  - acadgild supergroup           0 2018-07-07 09:26 /pigoutput
drwxr-xr-x  - acadgild supergroup           0 2018-07-06 19:00 /sqoopout
drwxr-xr-x  - acadgild supergroup           0 2018-07-06 20:10 /sqoopoutbyid
-rw-r--r--  1 acadgild supergroup          252 2018-07-06 16:16 /test.txt
drwxrwx---  - acadgild supergroup           0 2018-07-07 06:15 /tmp
drwxr-xr-x  - acadgild supergroup           0 2018-07-06 07:38 /user
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost pig]$
```

Step:2 : - Check the contain of test.txt

```
[acadgild@localhost pig]$ hadoop fs -cat /test.txt
18/07/07 14:01:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library
    tlin-java classes where applicable
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
Hi I am Manish Anand
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost pig]$
```

Now we will be perform wordmedian, wordmean and wordstandarddeviation operation on text.txt file using map reduce example jar which is present in VM.

## 1. WordMean

```
WordMean: A map/reduce program that counts the standard deviation of the length of the words in th
e input files.
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost pig]$ hadoop jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar word
mean /test.txt /output
18/07/07 12:47:04 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using bui
    tlin-java classes where applicable
18/07/07 12:47:05 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/07 12:47:07 INFO input.FileInputFormat: Total input paths to process : 1
```

```
Bytes Read=252
File Output Format Counters
Bytes Written=20
The mean is: 3.2
You have new mail in /var/spool/mail/acadgild
```

```
[acadgild@localhost pig]$ hadoop fs -ls /output
18/07/07 13:43:30 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using bui
    tlin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2018-07-07 12:47 /output/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 20 2018-07-07 12:47 /output/part-r-00000
[acadgild@localhost pig]$ hadoop fs -cat /output/part-r-00000
18/07/07 13:43:54 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using bui
    tlin-java classes where applicable
count 60
length 192
```

## 2.WordMedian

```
[acadgild@localhost pig]$  
[acadgild@localhost pig]$ hadoop jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar word  
median /test.txt /output  
18/07/07 13:48:48 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using bui  
ltin-java classes where applicable  
18/07/07 13:48:51 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032  
18/07/07 13:48:53 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement  
the Tool interface and execute your application with ToolRunner to remedy this.  
18/07/07 13:48:54 INFO input.FileInputFormat: Total input paths to process : 1  
18/07/07 13:48:54 INFO mapreduce.JobSubmitter: number of splits:1  
18/07/07 13:48:55 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530834392057_0032  
18/07/07 13:48:56 INFO impl.YarnClientImpl: Submitted application application_1530834392057_0032  
18/07/07 13:48:56 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_153083439  
2057_0032/  
18/07/07 13:48:56 INFO mapreduce.Job: Running job: job_1530834392057_0032  
18/07/07 13:49:16 INFO mapreduce.Job: Job job_1530834392057_0032 running in uber mode : false  
18/07/07 13:49:16 INFO mapreduce.Job: map 0% reduce 0%  
18/07/07 13:49:26 INFO mapreduce.Job: map 100% reduce 0%  
18/07/07 13:49:36 INFO mapreduce.Job: map 100% reduce 100%  
18/07/07 13:49:37 INFO mapreduce.Job: Job job_1530834392057_0032 completed successfully  
18/07/07 13:49:38 INFO mapreduce.Job: Counters: 49  
File System Counters  
FILE: Number of bytes read=46  
FILE: Number of bytes written=215129
```

```
GC time elapsed (ms)=178  
CPU time spent (ms)=2160  
Physical memory (bytes) snapshot=321417216  
Virtual memory (bytes) snapshot=4117938176  
Total committed heap usage (bytes)=222429184  
  
Shuffle Errors  
BAD_ID=0  
CONNECTION=0  
IO_ERROR=0  
WRONG_LENGTH=0  
WRONG_MAP=0  
WRONG_REDUCE=0  
  
File Input Format Counters  
Bytes Read=252  
File Output Format Counters  
Bytes Written=20  
  
The median is: 2  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost pig]$
```

```
WRONG_LENGTH=0  
WRONG_MAP=0  
WRONG_REDUCE=0  
File Input Format Counters  
Bytes Read=252  
File Output Format Counters  
Bytes Written=20  
  
The median is: 2  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost pig]$ hadoop fs -cat /output/part-r-00000  
18/07/07 13:51:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...  
ltin-java classes where applicable  
1 12  
2 24  
5 12  
6 12  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost pig]$
```



### 3. WordStandardDeviation

```
MultiExec Tunneling Packages Settings Help X server Exit
6. 127.0.0.1 (acadgild) 7. 127.0.0.1 (acadgild) 8. 127.0.0.1 (acadgild)
Deleted /output
[acadgild@localhost pig]$ hadoop jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar word
standarddeviation /test.txt /output
18/07/07 13:55:56 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using bui
ltin-java classes where applicable
18/07/07 13:55:59 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/07 13:56:02 INFO input.FileInputFormat: Total input paths to process : 1
18/07/07 13:56:02 INFO mapreduce.JobSubmitter: number of splits:1
18/07/07 13:56:03 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530834392057_0033
18/07/07 13:56:04 INFO impl.YarnClientImpl: Submitted application application_1530834392057_0033
18/07/07 13:56:04 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_153083439
2057_0033/
18/07/07 13:56:04 INFO mapreduce.Job: Running job: job_1530834392057_0033
18/07/07 13:56:22 INFO mapreduce.Job: Job job_1530834392057_0033 running in uber mode : false
18/07/07 13:56:22 INFO mapreduce.Job: map 0% reduce 0%
18/07/07 13:56:35 INFO mapreduce.Job: map 100% reduce 0%
18/07/07 13:56:46 INFO mapreduce.Job: map 100% reduce 100%
18/07/07 13:56:47 INFO mapreduce.Job: Job job_1530834392057_0033 completed successfully
18/07/07 13:56:48 INFO mapreduce.Job: Counters: 49
File System Counters
  FILE: Number of bytes read=56
  FILE: Number of bytes written=215573
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=347
  HDFS: Number of bytes written=31
  HDFS: Number of read operations=6
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counters
  Launched map tasks=1
  Launched reduce tasks=1

Physical memory (bytes) snapshot=324616192
Virtual memory (bytes) snapshot=4122288128
Total committed heap usage (bytes)=222429184

Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=252
File Output Format Counters
  Bytes Written=31
The standard deviation is: 1.939071942966531
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost pig]$
```

```
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=252
File Output Format Counters
  Bytes Written=31
The standard deviation is: 1.939071942966531
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost pig]$ hadoop fs -cat /output/part-r-00000
18/07/07 14:06:52 WARN util.NativeCodeLoader: Unable to load native-hadoop library
from java classes where applicable
count      60
length     192
square     840
[acadgild@localhost pig]$
```