Task 1:
Execute (i) **WordMedian** , (ii) **WordMean** , (iii) **WordStandardDeviation** programs using hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.

(i) COMMAND:-

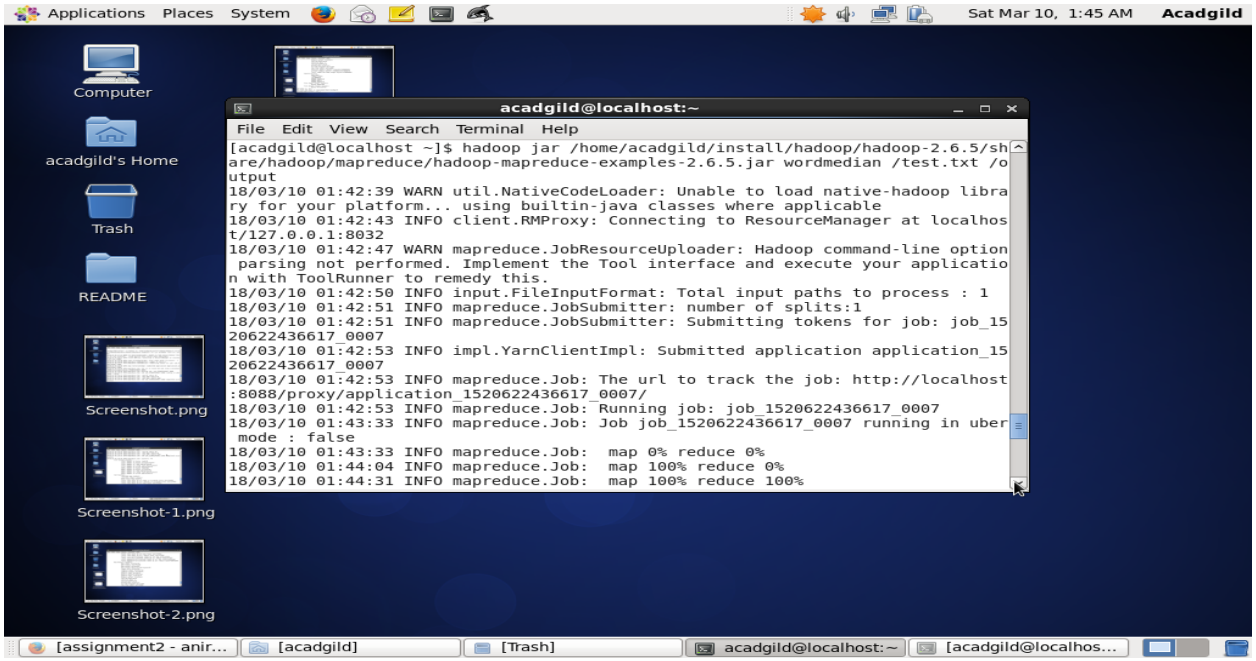
hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmedian /test.txt /output

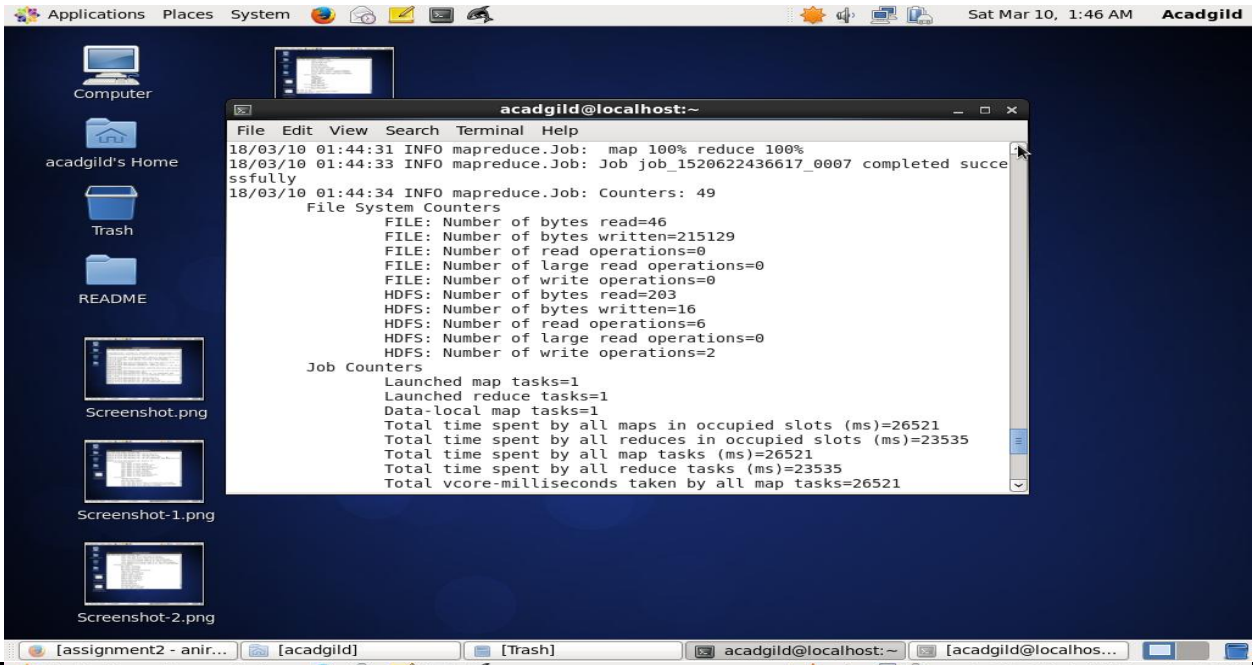
note:-At Acadgild VM,Here I have hadoop-mapreduce-examples-2.6.5.jar instead of hadoop-mapreduce-examples-2.9.0.jar .

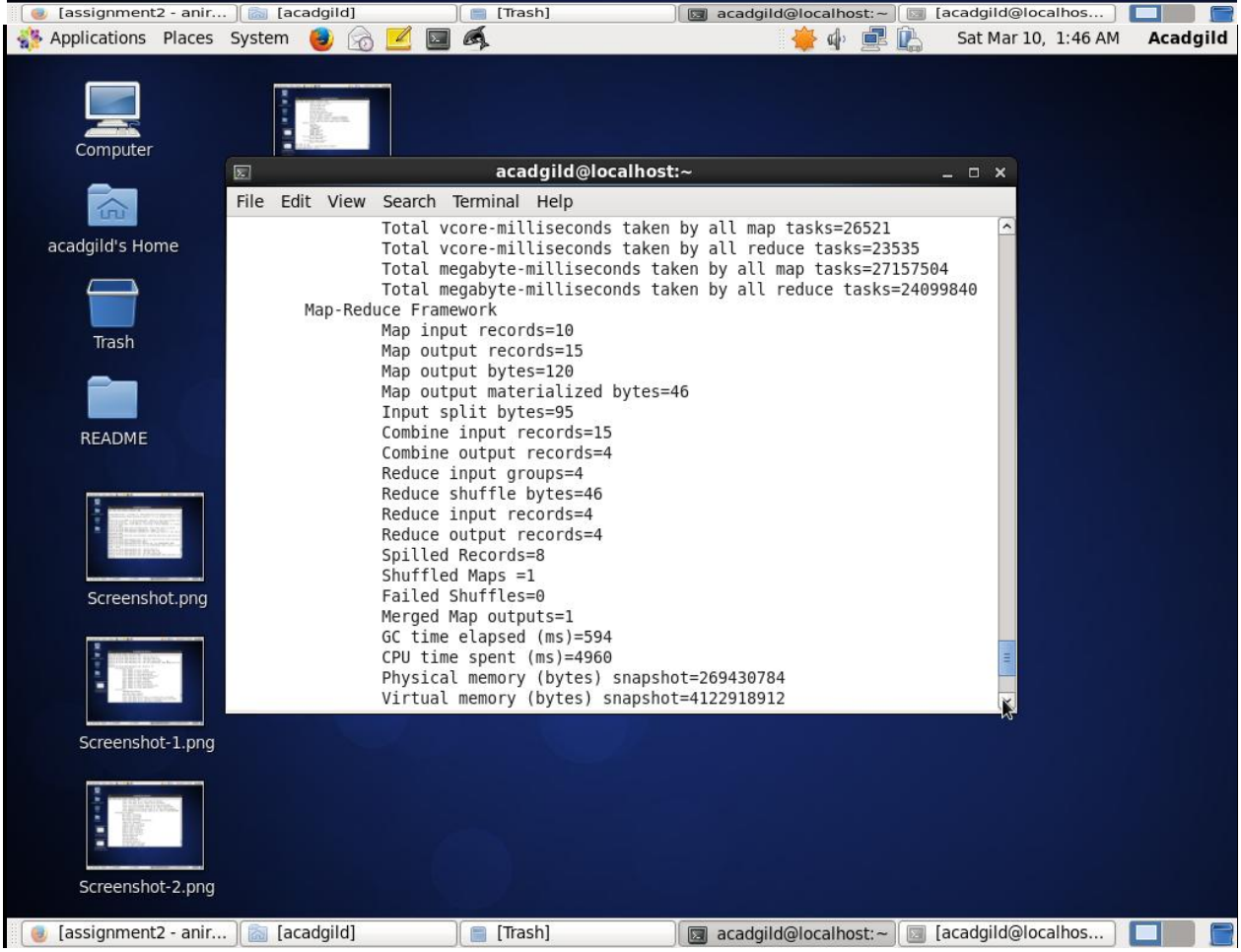
EXPLANATION:-

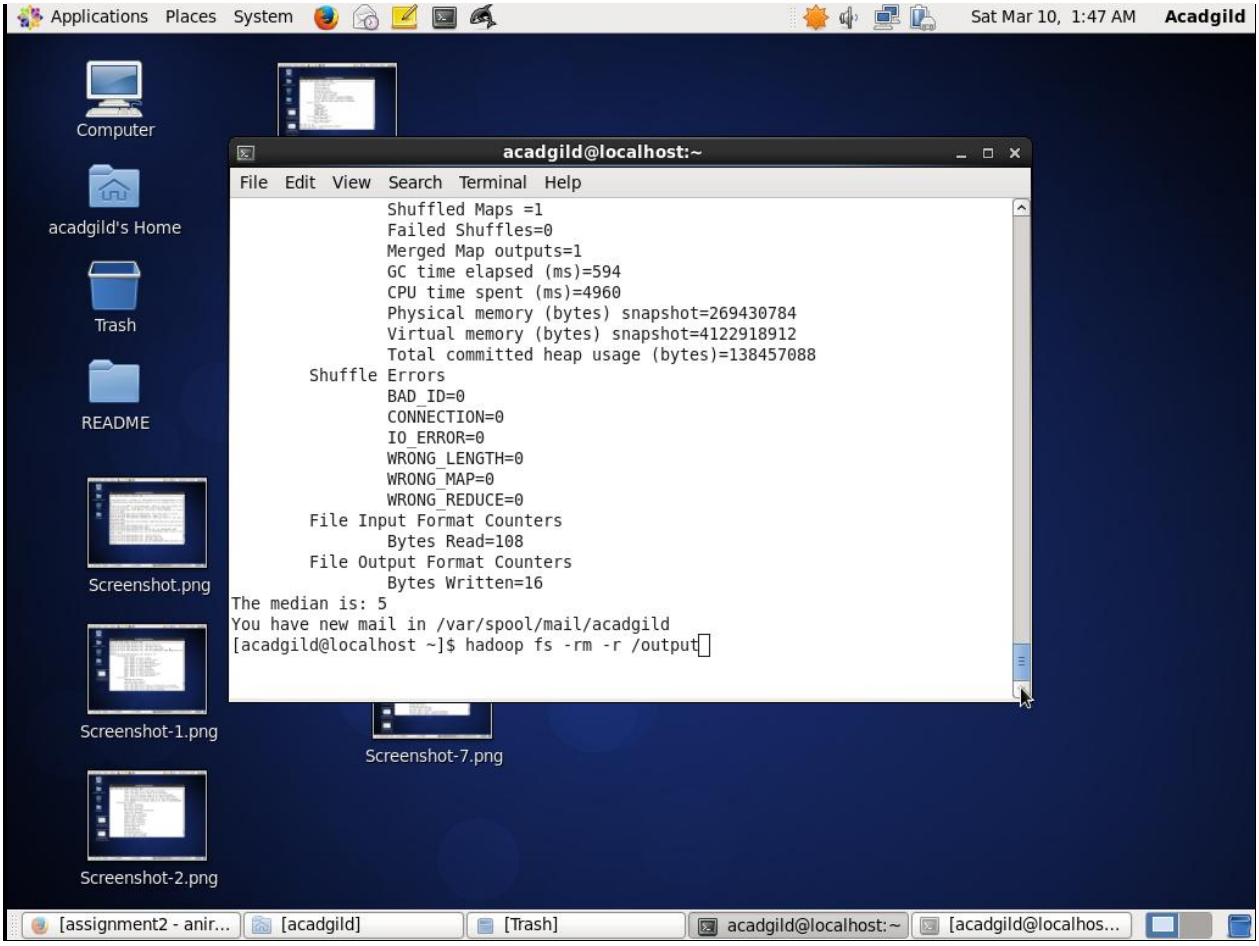
wordmedian:- the map/reduce program that counts the median length of the words in the input files.

OUTPUT:-




Applications Places System
Sat Mar 10, 1:46 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
Screenshot-1.png
Screenshot-2.png
acadgild@localhost:~
File Edit View Search Terminal Help
18/03/10 01:44:31 INFO mapreduce.Job: map 100% reduce 100%
18/03/10 01:44:33 INFO mapreduce.Job: Job job_1520622436617_0007 completed succe
ssfully
18/03/10 01:44:34 INFO mapreduce.Job: Counters: 49
File System Counters
FILE: Number of bytes read=46
FILE: Number of bytes written=215129
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=203
HDFS: Number of bytes written=16
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
Data-local map tasks=1
Total time spent by all maps in occupied slots (ms)=26521
Total time spent by all reduces in occupied slots (ms)=23535
Total time spent by all map tasks (ms)=26521
Total time spent by all reduce tasks (ms)=23535
Total vcore-milliseconds taken by all map tasks=26521
[assignment2 - anir...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...


Applications Places System
Sat Mar 10, 1:46 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
Screenshot-1.png
Screenshot-2.png
acadgild@localhost:~
File Edit View Search Terminal Help
Total vcore-milliseconds taken by all map tasks=26521
Total vcore-milliseconds taken by all reduce tasks=23535
Total megabyte-milliseconds taken by all map tasks=27157504
Total megabyte-milliseconds taken by all reduce tasks=24099840
Map-Reduce Framework
Map input records=10
Map output records=15
Map output bytes=120
Map output materialized bytes=46
Input split bytes=95
Combine input records=15
Combine output records=4
Reduce input groups=4
Reduce shuffle bytes=46
Reduce input records=4
Reduce output records=4
Spilled Records=8
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=594
CPU time spent (ms)=4960
Physical memory (bytes) snapshot=269430784
Virtual memory (bytes) snapshot=4122918912
[assignment2 - anir...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...



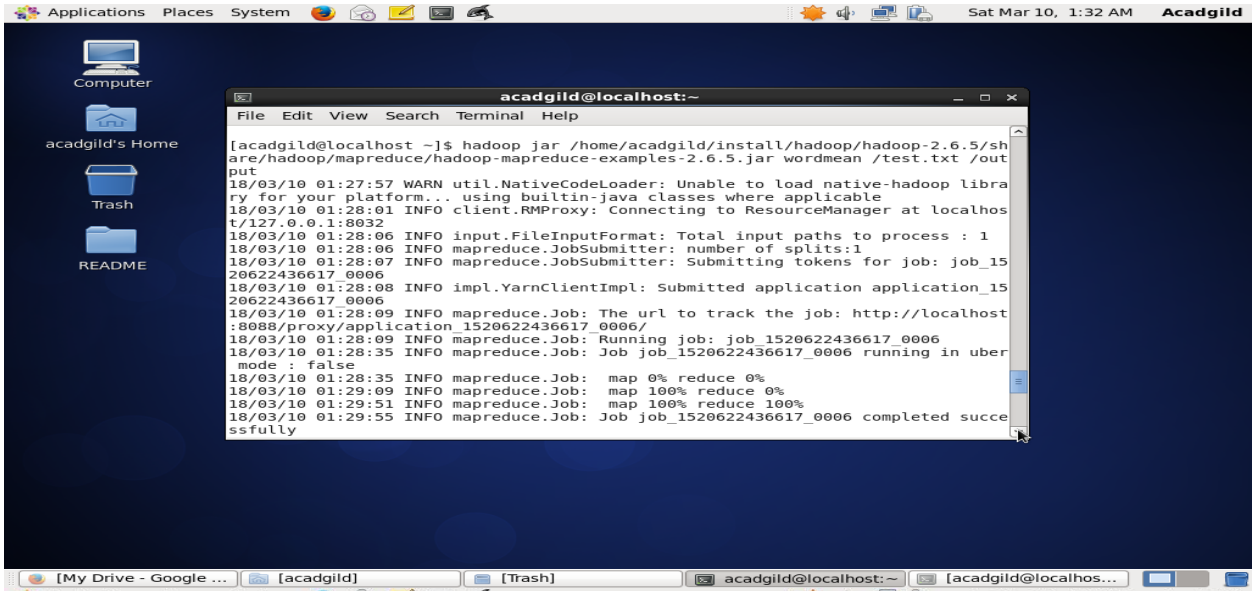
(ii) COMMAND:-

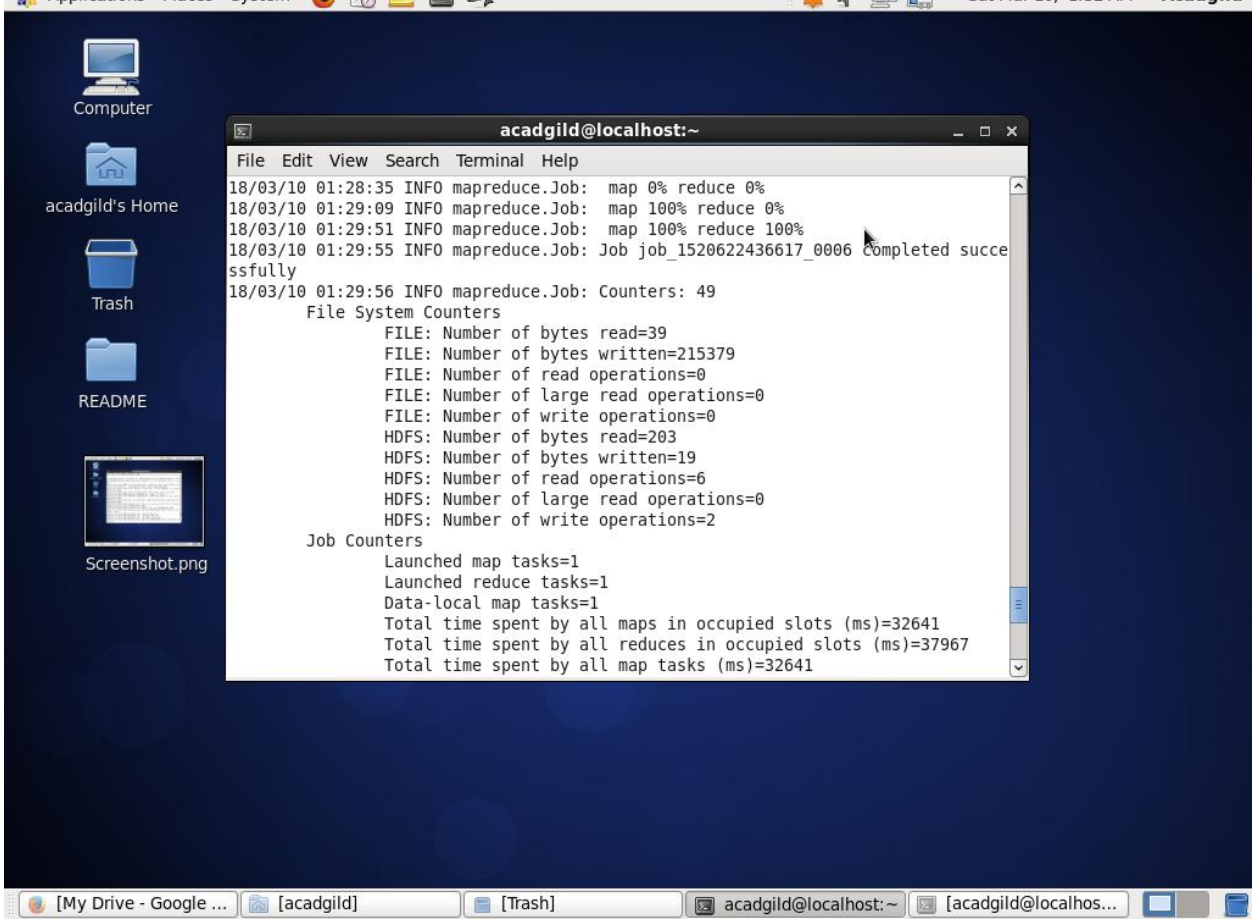
hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmean /test.txt /output

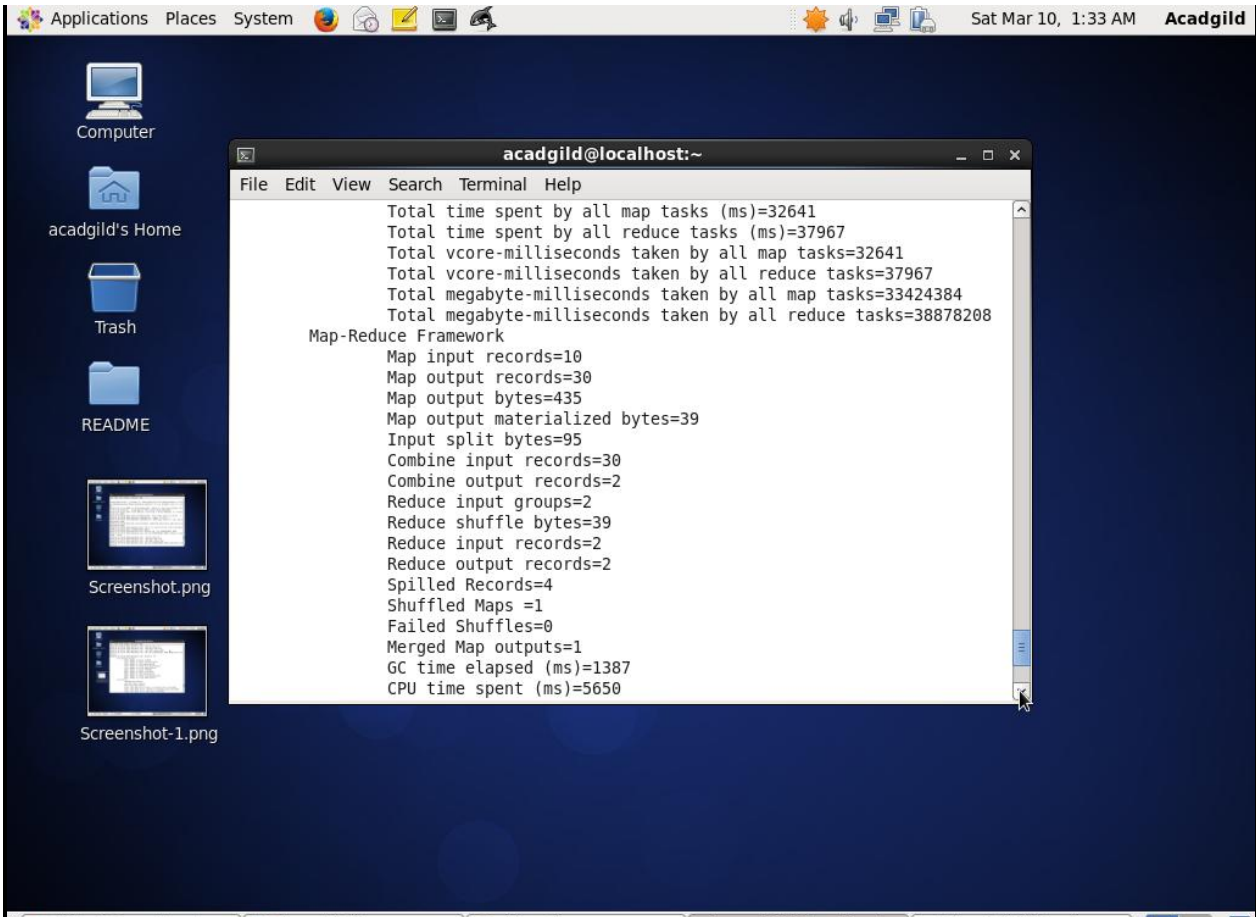
EXPLANATION:-

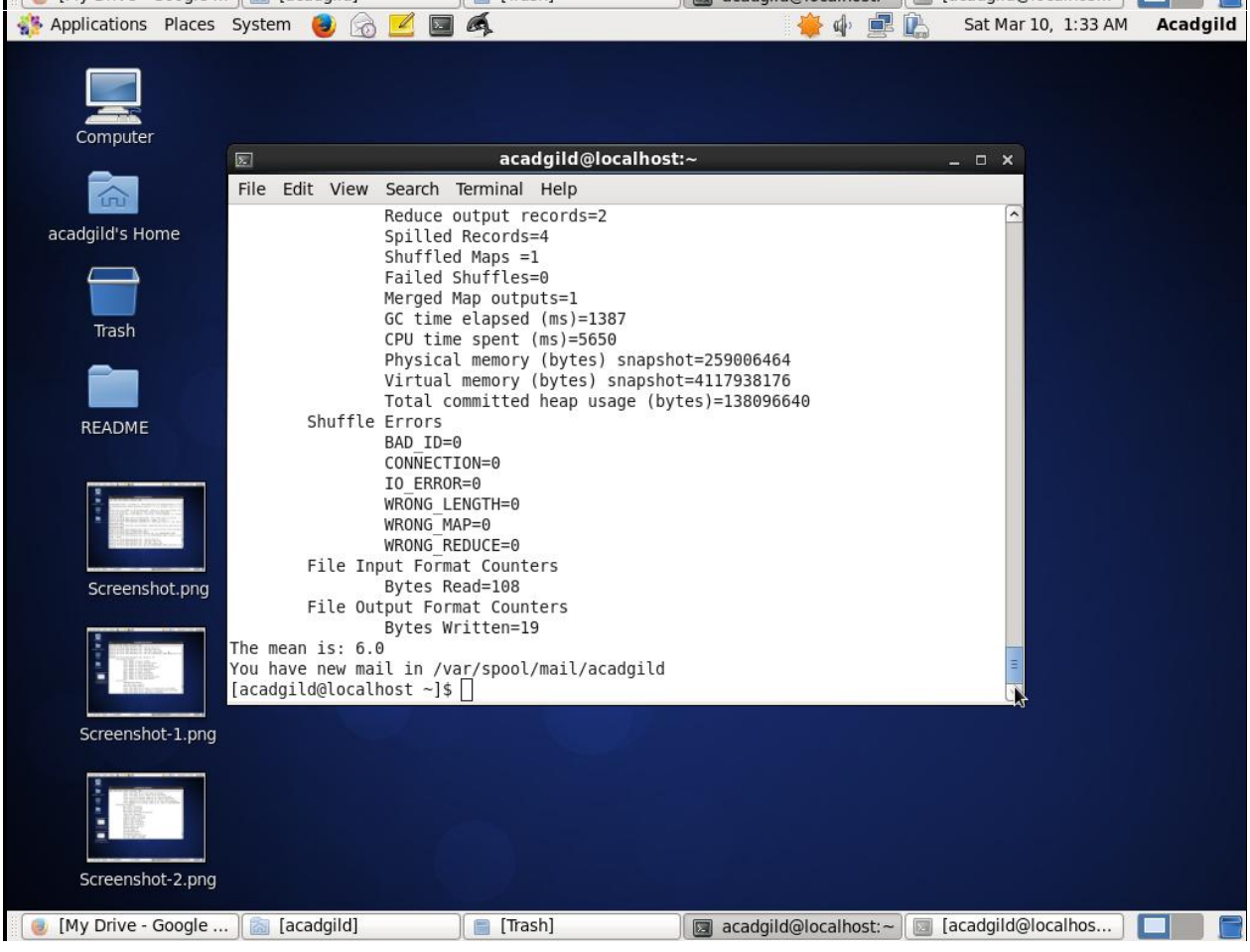
wordmean:- the map/reduce program that counts the average length of the words in the input files.

OUTPUT:-


Applications Places System
Sat Mar 10, 1:32 AM Acadgild
Computer
acadgild's Home
Trash
README
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmean /test.txt /output
18/03/10 01:27:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/03/10 01:28:01 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/03/10 01:28:06 INFO input.FileInputFormat: Total input paths to process : 1
18/03/10 01:28:06 INFO mapreduce.JobSubmitter: number of splits:1
18/03/10 01:28:07 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1520622436617_0006
18/03/10 01:28:08 INFO impl.YarnClientImpl: Submitted application application_1520622436617_0006
18/03/10 01:28:09 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1520622436617_0006/
18/03/10 01:28:09 INFO mapreduce.Job: Running job: job_1520622436617_0006
18/03/10 01:28:35 INFO mapreduce.Job: Job job_1520622436617_0006 running in uber mode : false
18/03/10 01:28:35 INFO mapreduce.Job: map 0% reduce 0%
18/03/10 01:29:09 INFO mapreduce.Job: map 100% reduce 0%
18/03/10 01:29:51 INFO mapreduce.Job: map 100% reduce 100%
18/03/10 01:29:55 INFO mapreduce.Job: Job job_1520622436617_0006 completed successfully
[My Drive - Google ...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...


Applications Places System
Sat Mar 10, 1:32 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
acadgild@localhost:~
File Edit View Search Terminal Help
18/03/10 01:28:35 INFO mapreduce.Job: map 0% reduce 0%
18/03/10 01:29:09 INFO mapreduce.Job: map 100% reduce 0%
18/03/10 01:29:51 INFO mapreduce.Job: map 100% reduce 100%
18/03/10 01:29:55 INFO mapreduce.Job: Job job_1520622436617_0006 completed successfully
18/03/10 01:29:56 INFO mapreduce.Job: Counters: 49
File System Counters
FILE: Number of bytes read=39
FILE: Number of bytes written=215379
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=203
HDFS: Number of bytes written=19
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
Data-local map tasks=1
Total time spent by all maps in occupied slots (ms)=32641
Total time spent by all reduces in occupied slots (ms)=37967
Total time spent by all map tasks (ms)=32641
[My Drive - Google ...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...


Applications Places System
Sat Mar 10, 1:33 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
Screenshot-1.png
acadgild@localhost:~
File Edit View Search Terminal Help
Total time spent by all map tasks (ms)=32641
Total time spent by all reduce tasks (ms)=37967
Total vcore-milliseconds taken by all map tasks=32641
Total vcore-milliseconds taken by all reduce tasks=37967
Total megabyte-milliseconds taken by all map tasks=33424384
Total megabyte-milliseconds taken by all reduce tasks=38878208
Map-Reduce Framework
Map input records=10
Map output records=30
Map output bytes=435
Map output materialized bytes=39
Input split bytes=95
Combine input records=30
Combine output records=2
Reduce input groups=2
Reduce shuffle bytes=39
Reduce input records=2
Reduce output records=2
Spilled Records=4
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=1387
CPU time spent (ms)=5650
[My Drive - Google ...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...


Applications Places System
Sat Mar 10, 1:33 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
Screenshot-1.png
Screenshot-2.png
acadgild@localhost:~
File Edit View Search Terminal Help
Reduce output records=2
Spilled Records=4
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=1387
CPU time spent (ms)=5650
Physical memory (bytes) snapshot=259006464
Virtual memory (bytes) snapshot=4117938176
Total committed heap usage (bytes)=138096640
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=108
File Output Format Counters
Bytes Written=19
The mean is: 6.0
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
[My Drive - Google ...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...

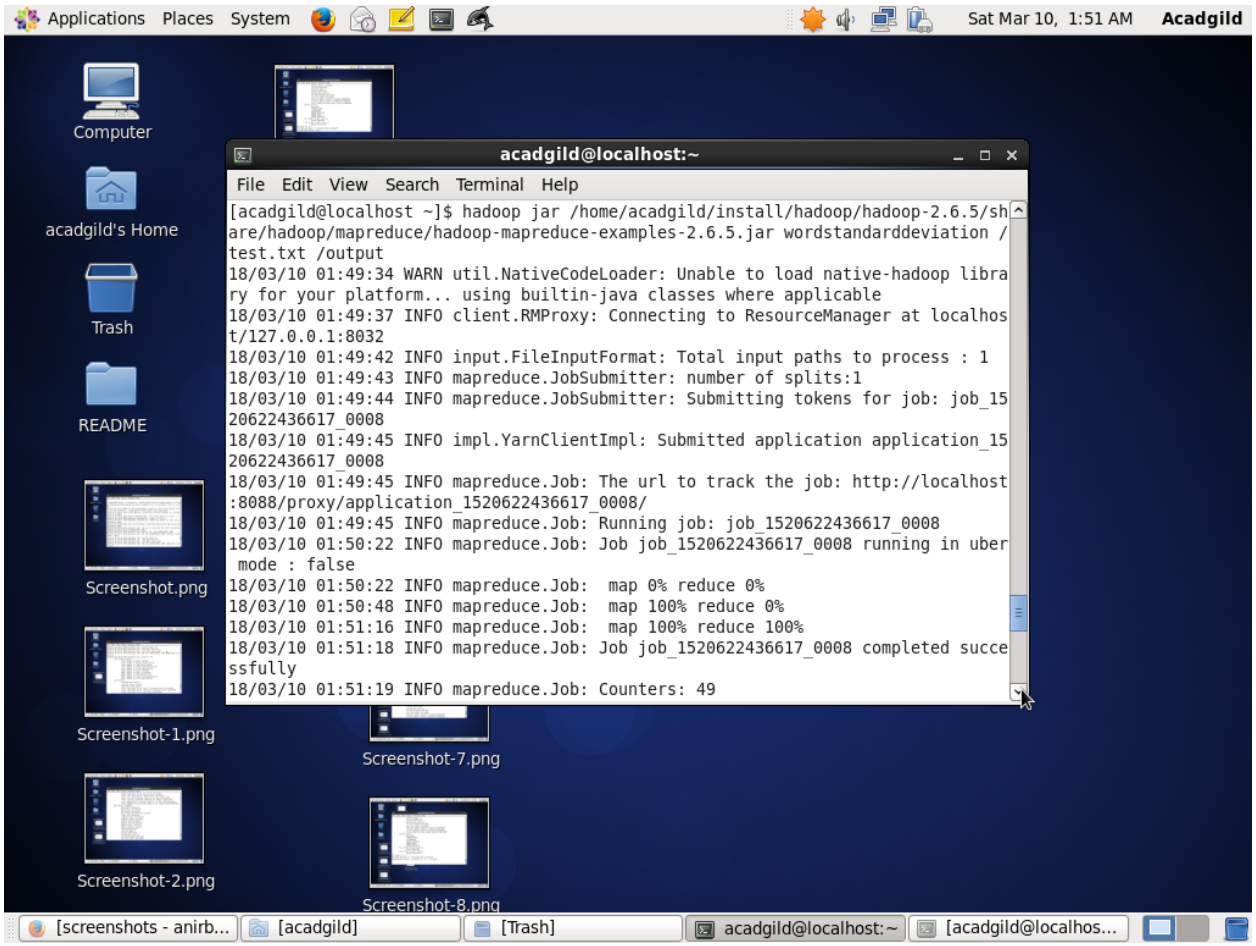
(iii) COMMAND:-

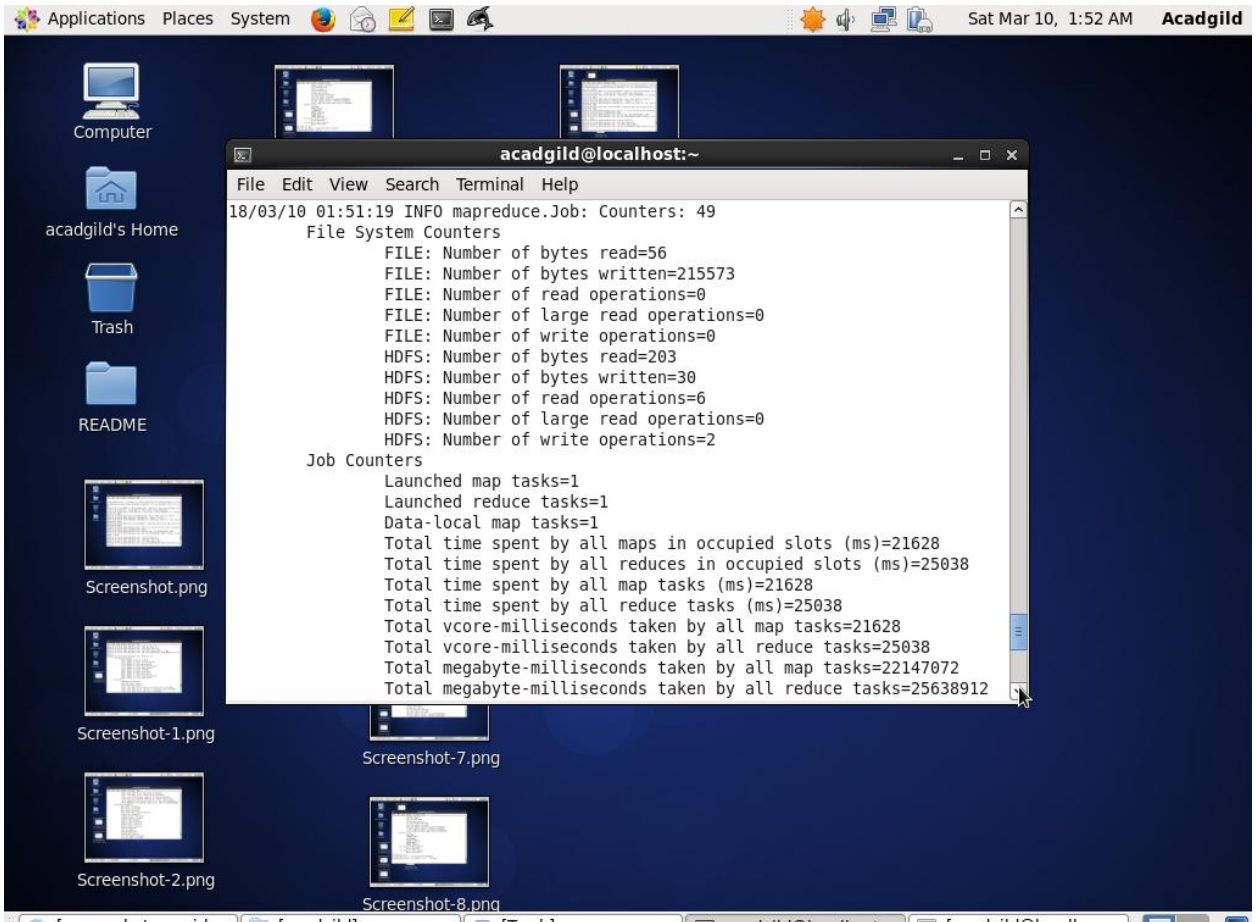
hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /test.txt /output

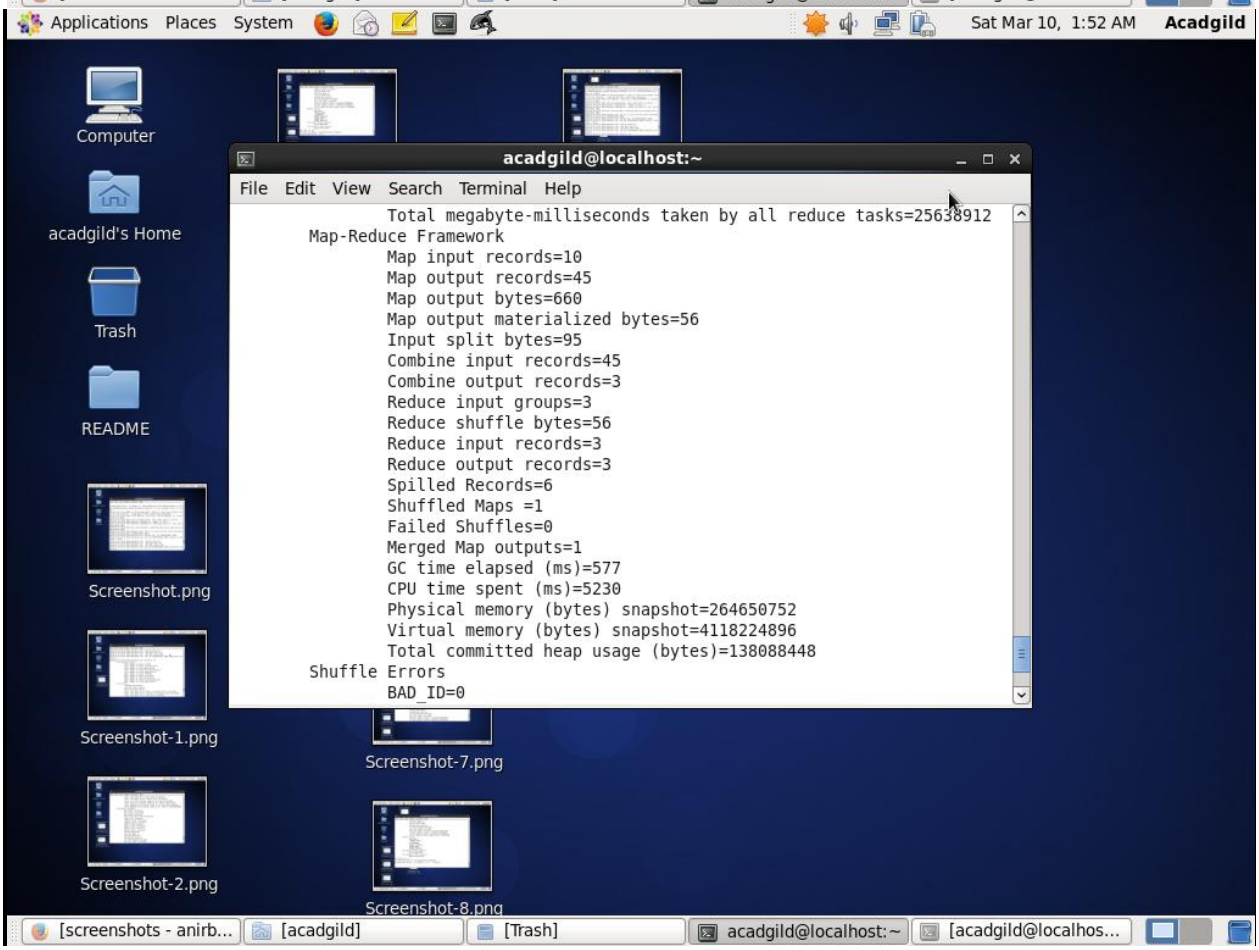
EXPLANATION:-

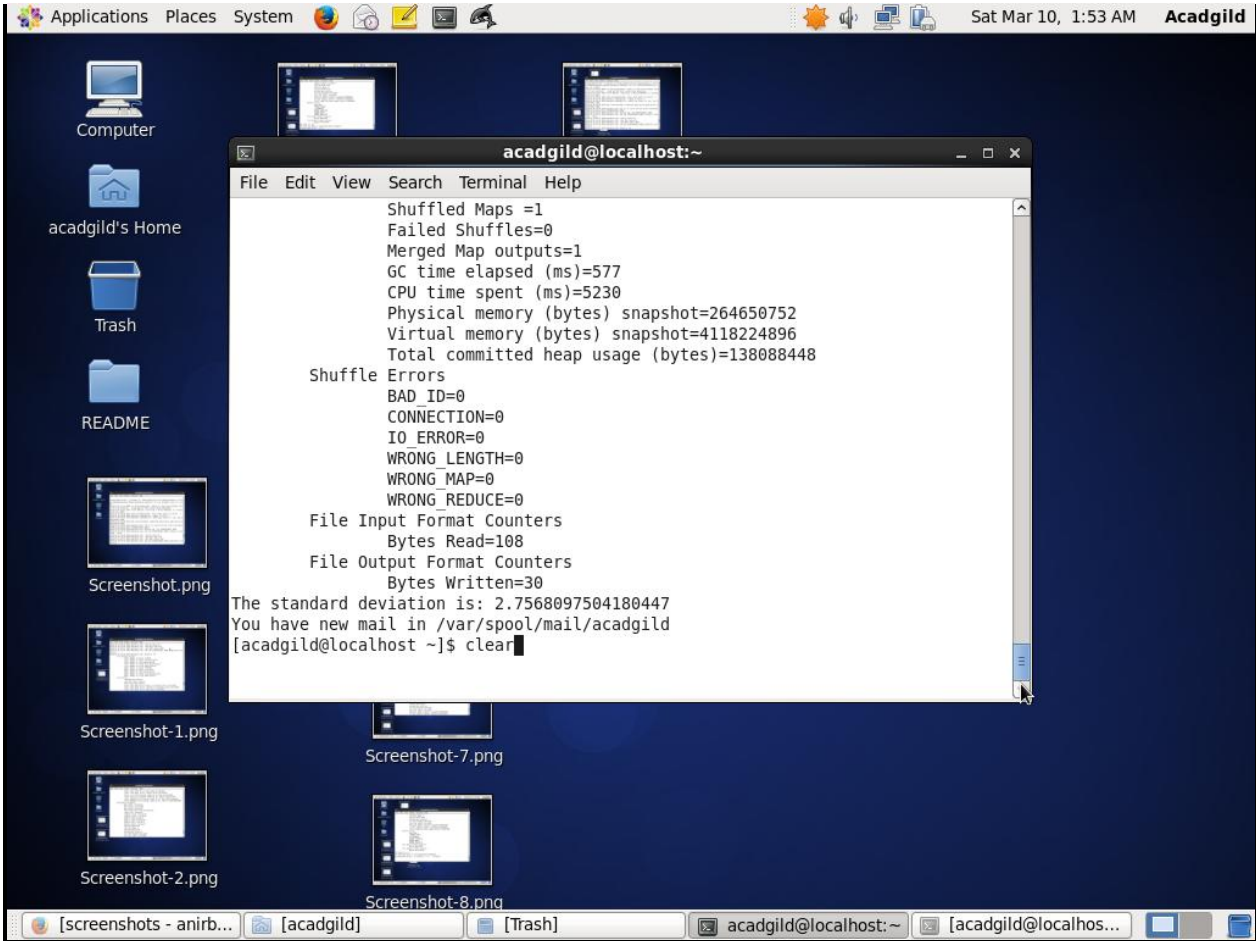
wordstandarddeviation:- the map/reduce program that counts the standard deviation of the length of the words in the input files.

OUTPUT:-




Applications Places System
Sat Mar 10, 1:52 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
Screenshot-1.png
Screenshot-2.png
Screenshot-7.png
Screenshot-8.png
acadgild@localhost:~
File Edit View Search Terminal Help
18/03/10 01:51:19 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=203
HDFS: Number of bytes written=30
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
Data-local map tasks=1
Total time spent by all maps in occupied slots (ms)=21628
Total time spent by all reduces in occupied slots (ms)=25038
Total time spent by all map tasks (ms)=21628
Total time spent by all reduce tasks (ms)=25038
Total vcore-milliseconds taken by all map tasks=21628
Total vcore-milliseconds taken by all reduce tasks=25038
Total megabyte-milliseconds taken by all map tasks=22147072
Total megabyte-milliseconds taken by all reduce tasks=25638912
[screenshots - anirb...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...


Applications Places System
Sat Mar 10, 1:52 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
Screenshot-1.png
Screenshot-2.png
Screenshot-7.png
Screenshot-8.png
acadgild@localhost:~
File Edit View Search Terminal Help
Total megabyte-milliseconds taken by all reduce tasks=25638912
Map-Reduce Framework
Map input records=10
Map output records=45
Map output bytes=660
Map output materialized bytes=56
Input split bytes=95
Combine input records=45
Combine output records=3
Reduce input groups=3
Reduce shuffle bytes=56
Reduce input records=3
Reduce output records=3
Spilled Records=6
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=577
CPU time spent (ms)=5230
Physical memory (bytes) snapshot=264650752
Virtual memory (bytes) snapshot=4118224896
Total committed heap usage (bytes)=138088448
Shuffle Errors
BAD_ID=0
[screenshots - anirb...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...


Applications Places System
Sat Mar 10, 1:53 AM Acadgild
Computer
acadgild's Home
Trash
README
Screenshot.png
Screenshot-1.png
Screenshot-2.png
Screenshot-7.png
Screenshot-8.png
acadgild@localhost:~
File Edit View Search Terminal Help
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=577
CPU time spent (ms)=5230
Physical memory (bytes) snapshot=264650752
Virtual memory (bytes) snapshot=4118224896
Total committed heap usage (bytes)=138088448
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=108
File Output Format Counters
Bytes Written=30
The standard deviation is: 2.7568097504180447
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ clear
[screenshots - anirb...
[acadgild]
[Trash]
acadgild@localhost:~
[acadgild@localhos...