```
                      • MobaXterm 10.4 •
            (SSH client, X-server and networking tools)

   → SSH session to acadgild@192.168.56.2
     • SSH compression : v
     • SSH-browser     : v
     • X11-forwarding  : v  (remote display is forwarded through SSH)
     • DISPLAY         : v  (automatically set on remote server)


   → For more info, ctrl+click on help or visit our website
```

Last login: Tue Jul 3 13:45:11 2018 from 192.168.56.1
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$ cd ..
[acadgild.mmisra home]$

Look for map reduce example jar
[acadgild.mmisra home]$ **ls -la**
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/
total 4888
drwxrwxr-x. 5 acadgild acadgild   4096 Oct 3 2016 .
drwxrwxr-x. 9 acadgild acadgild   4096 Oct 3 2016 ..
-rw-rw-r--. 1 acadgild acadgild 526732 Oct 3 2016
hadoop-mapreduce-client-app-2.6.5.jar
-rw-rw-r--. 1 acadgild acadgild 686773 Oct 3 2016
hadoop-mapreduce-client-common-2.6.5.jar
-rw-rw-r--. 1 acadgild acadgild 1535776 Oct 3 2016
hadoop-mapreduce-client-core-2.6.5.jar
-rw-rw-r--. 1 acadgild acadgild 259326 Oct 3 2016 hadoop-mapreduce-client-hs-2.6.5.jar
-rw-rw-r--. 1 acadgild acadgild   27489 Oct 3 2016
hadoop-mapreduce-client-hs-plugins-2.6.5.jar
-rw-rw-r--. 1 acadgild acadgild   61309 Oct 3 2016
hadoop-mapreduce-client-jobclient-2.6.5.jar
-rw-rw-r--. 1 acadgild acadgild 1514166 Oct 3 2016
hadoop-mapreduce-client-jobclient-2.6.5-tests.jar
-rw-rw-r--. 1 acadgild acadgild   67762 Oct 3 2016
hadoop-mapreduce-client-shuffle-2.6.5.jar
-rw-rw-r--. 1 acadgild acadgild 292710 Oct 3 2016 **hadoop-mapreduce-examples-2.6.5.jar**
drwxrwxr-x. 2 acadgild acadgild   4096 Oct 3 2016 lib
drwxrwxr-x. 2 acadgild acadgild   4096 Oct 3 2016 lib-examples
drwxrwxr-x. 2 acadgild acadgild   4096 Oct 3 2016 sources

Run the example JAR to see command line options
[acadgild.mmisra home]$
[acadgild.mmisra home]$
[acadgild.mmisra home]$ **hadoop jar**
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa
mples-2.6.5.jar
An example program must be given as the first argument.
Valid program names are:
  aggregatewordcount: An Aggregate based map/reduce program that counts the words in the
input files.
  aggregatewordhist: An Aggregate based map/reduce program that computes the histogram of
the words in the input files.
  bbp: A map/reduce program that uses Bailey-Borwein-Plouffe to compute exact digits of
Pi.
  dbcount: An example job that count the pageview counts from a database.
  distbbp: A map/reduce program that uses a BBP-type formula to compute exact bits of Pi.
  grep: A map/reduce program that counts the matches of a regex in the input.
  join: A job that effects a join over sorted, equally partitioned datasets

```
  multifilewc: A job that counts words from several files.
  pentomino: A map/reduce tile laying program to find solutions to pentomino problems.
  pi: A map/reduce program that estimates Pi using a quasi-Monte Carlo method.
  randomtextwriter: A map/reduce program that writes 10GB of random textual data per node.
  randomwriter: A map/reduce program that writes 10GB of random data per node.
  secondarysort: An example defining a secondary sort to the reduce.
  sort: A map/reduce program that sorts the data written by the random writer.
  sudoku: A sudoku solver.
  teragen: Generate data for the terasort
  terasort: Run the terasort
  teravalidate: Checking results of terasort
  wordcount: A map/reduce program that counts the words in the input files.
  wordmean: A map/reduce program that counts the average length of the words in the input
files.
  wordmedian: A map/reduce program that counts the median length of the words in the input
files.
  wordstandarddeviation: A map/reduce program that counts the standard deviation of the
length of the words in the input files.
[acadgild.mmisra home]$
[acadgild.mmisra home]$
[acadgild.mmisra home]$
[acadgild.mmisra home]$ clear
[acadgild.mmisra home]$
```

**# Prepare the input file which is put in HDFS already called file327.txt which is around
327MB file**

```
[acadgild.mmisra home]$ hadoop fs -ls /files
18/07/03 15:15:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup  327155712 2018-06-24 10:51 /files/file327.txt
-rw-r--r--   1 acadgild supergroup        336 2018-06-24 10:51 /files/test.txt
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra home]$
[acadgild.mmisra home]$
```

 # Check to make sure the desired output directory does not exist on HDFS

```
[acadgild.mmisra home]$ hadoop fs -ls /
18/07/03 15:15:33 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 7 items
drwxr-xr-x   - acadgild supergroup          0 2018-06-24 10:51 /files
drwxr-xr-x   - acadgild supergroup          0 2018-06-24 10:50 /mohit
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 09:03 /out1
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 10:01 /sqoopout
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 11:21 /sqoopoutbyid
drwx------   - acadgild supergroup          0 2018-06-24 11:27 /tmp
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 14:12 /user
[acadgild.mmisra home]$
[acadgild.mmisra home]$
```

**# Run the wordmean option which is a map/reduce program that counts the average length of
the words in the input files. Here the input file is /files/file327.txt and the output result
would be stored in /wordmean_out directory**

```
6.5.jar wordmean home]$ hadoop jar
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa
mples-2.6.jar wordmean
Usage: wordmean <in> <out>
[acadgild.mmisra home]$
[acadgild.mmisra home]$
```

[acadgild.mmisra home]$ **hadoop jar**
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa
mples-2.6.jar wordmean /files/file327.txt /wordmean_out

18/07/03 15:16:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
18/07/03 15:16:24 INFO client.RMProxy: Connecting to ResourceManager at
localhost/127.0.0.1:8032

 # only one input file to process
18/07/03 15:16:25 INFO input.FileInputFormat: Total input paths to process : 1

 # There are 3 input splits as the file size is 327MB and each HDFS block size is 128MB.
First 2 block would be of 128MB and third block would be 71MB. Therefore 3 mappers will
be created

18/07/03 15:16:25 INFO mapreduce.JobSubmitter: number of splits:3
18/07/03 15:16:26 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1530595889847_0001
18/07/03 15:16:26 INFO impl.YarnClientImpl: Submitted application
application_1530595889847_0001
18/07/03 15:16:26 INFO mapreduce.Job: The url to track the job:
 # URL to track the job

http://localhost:8088/proxy/application_1530595889847_0001/

18/07/03 15:16:26 INFO mapreduce.Job: Running job: job_1530595889847_0001
18/07/03 15:16:35 INFO mapreduce.Job: Job job_1530595889847_0001 running in uber mode :
false
18/07/03 15:16:35 INFO mapreduce.Job: map 0% reduce 0%
18/07/03 15:16:50 INFO mapreduce.Job: map 3% reduce 0%
18/07/03 15:16:52 INFO mapreduce.Job: map 6% reduce 0%
18/07/03 15:16:53 INFO mapreduce.Job: map 8% reduce 0%
18/07/03 15:16:55 INFO mapreduce.Job: map 10% reduce 0%
18/07/03 15:16:56 INFO mapreduce.Job: map 11% reduce 0%
18/07/03 15:16:58 INFO mapreduce.Job: map 14% reduce 0%
18/07/03 15:16:59 INFO mapreduce.Job: map 16% reduce 0%
18/07/03 15:17:01 INFO mapreduce.Job: map 18% reduce 0%
18/07/03 15:17:02 INFO mapreduce.Job: map 21% reduce 0%
18/07/03 15:17:04 INFO mapreduce.Job: map 24% reduce 0%
18/07/03 15:17:05 INFO mapreduce.Job: map 26% reduce 0%
18/07/03 15:17:07 INFO mapreduce.Job: map 29% reduce 0%
18/07/03 15:17:08 INFO mapreduce.Job: map 32% reduce 0%
18/07/03 15:17:10 INFO mapreduce.Job: map 34% reduce 0%
18/07/03 15:17:11 INFO mapreduce.Job: map 37% reduce 0%
18/07/03 15:17:13 INFO mapreduce.Job: map 40% reduce 0%
18/07/03 15:17:14 INFO mapreduce.Job: map 43% reduce 0%
18/07/03 15:17:16 INFO mapreduce.Job: map 58% reduce 0%
18/07/03 15:17:19 INFO mapreduce.Job: map 61% reduce 0%
18/07/03 15:17:22 INFO mapreduce.Job: map 63% reduce 0%
18/07/03 15:17:25 INFO mapreduce.Job: map 64% reduce 0%
18/07/03 15:17:28 INFO mapreduce.Job: map 66% reduce 0%
18/07/03 15:17:31 INFO mapreduce.Job: map 68% reduce 0%
18/07/03 15:17:34 INFO mapreduce.Job: map 70% reduce 11%
18/07/03 15:17:37 INFO mapreduce.Job: map 72% reduce 11%
18/07/03 15:17:38 INFO mapreduce.Job: map 84% reduce 11%
18/07/03 15:17:40 INFO mapreduce.Job: map 84% reduce 22%
18/07/03 15:17:41 INFO mapreduce.Job: map 86% reduce 22%
18/07/03 15:17:44 INFO mapreduce.Job: map 87% reduce 22%
18/07/03 15:17:47 INFO mapreduce.Job: map 100% reduce 22%
18/07/03 15:17:48 INFO mapreduce.Job: map 100% reduce 100%

 # Job completed successfully

18/07/03 15:17:49 INFO mapreduce.Job: Job job_1530595889847_0001 completed successfully
```
18/07/03 15:17:49 INFO mapreduce.Job: Counters: 50
        File System Counters
                FILE: Number of bytes read=2976
                FILE: Number of bytes written=433794
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=327164216
                HDFS: Number of bytes written=32
                HDFS: Number of read operations=12
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
```
 # somehow one map task was killed.
```
                Killed map tasks=1
                Launched map tasks=4
```
 # number of reducers is one
```
                Launched reduce tasks=1
                Data-local map tasks=4
                Total time spent by all maps in occupied slots (ms)=198322
                Total time spent by all reduces in occupied slots (ms)=29130
                Total time spent by all map tasks (ms)=198322
                Total time spent by all reduce tasks (ms)=29130
                Total vcore-milliseconds taken by all map tasks=198322
                Total vcore-milliseconds taken by all reduce tasks=29130
                Total megabyte-milliseconds taken by all map tasks=203081728
                Total megabyte-milliseconds taken by all reduce tasks=29829120
        Map-Reduce Framework
                Map input records=8388608
                Map output records=134217728
                Map output bytes=1946157056
                Map output materialized bytes=117
                Input split bytes=312
                Combine input records=134217830
                Combine output records=108
                Reduce input groups=2
                Reduce shuffle bytes=117
                Reduce input records=6
                Reduce output records=2
                Spilled Records=166
                Shuffled Maps =3
                Failed Shuffles=0
                Merged Map outputs=3
                GC time elapsed (ms)=1284
                CPU time spent (ms)=89900
                Physical memory (bytes) snapshot=945872896
                Virtual memory (bytes) snapshot=8333389824
                Total committed heap usage (bytes)=738197504
        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=327163904
        File Output Format Counters
                Bytes Written=32
```

 # The result – means length of all the words in the file.

The mean is: 3.875
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra home]$

 # check for job (reducer output)
[acadgild.mmisra home]$ hadoop fs -ls /wordmean_out
18/07/03 15:18:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where ap   plicable
Found 2 items
-rw-r--r--   1 acadgild supergroup   0 2018-07-03 15:17 /wordmean_out/_SUCCESS
-rw-r--r--   1 acadgild supergroup   32 2018-07-03 15:17 /wordmean_out/part-r-00000
[acadgild.mmisra home]$ hadoop fs -cat /wordmean_out/part-r-00000
18/07/03 15:18:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
count   67108864
length 260046848

#check for correctness, total length of all the words divided by the count of words
Mean = length/count = 3.875

[acadgild.mmisra home]$ **hadoop jar**
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa
mples-2.6.jar

An example program must be given as the first argument.
Valid program names are:
  aggregatewordcount: An Aggregate based map/reduce program that counts the words in the
input files.
  aggregatewordhist: An Aggregate based map/reduce program that computes the histogram of
the words in the input files.
  bbp: A map/reduce program that uses Bailey-Borwein-Plouffe to compute exact digits of
Pi.
  dbcount: An example job that count the pageview counts from a database.
  distbbp: A map/reduce program that uses a BBP-type formula to compute exact bits of Pi.
  grep: A map/reduce program that counts the matches of a regex in the input.
  join: A job that effects a join over sorted, equally partitioned datasets
  multifilewc: A job that counts words from several files.
  pentomino: A map/reduce tile laying program to find solutions to pentomino problems.
  pi: A map/reduce program that estimates Pi using a quasi-Monte Carlo method.
  randomtextwriter: A map/reduce program that writes 10GB of random textual data per node.
  randomwriter: A map/reduce program that writes 10GB of random data per node.
  secondarysort: An example defining a secondary sort to the reduce.
  sort: A map/reduce program that sorts the data written by the random writer.
  sudoku: A sudoku solver.
  teragen: Generate data for the terasort
  terasort: Run the terasort
  teravalidate: Checking results of terasort
  wordcount: A map/reduce program that counts the words in the input files.
  wordmean: A map/reduce program that counts the average length of the words in the input
files.
  wordmedian: A map/reduce program that counts the median length of the words in the input
files.
  wordstandarddeviation: A map/reduce program that counts the standard deviation of the
length of the words in the input files.
You have new mail in /var/spool/mail/acadgild

 # Run the wordmedian option which is a map/reduce program that counts the median length
of the words in the input files. We run the program on the same file327.txt and the output
is stored on HDFS under /wordmedian_out directory

```
[acadgild.mmisra home]$ hadoop jar
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa
mples-2.6 wordmeadian
Usage: wordmedian <in> <out>
[acadgild.mmisra home]$ hadoop jar
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa
mples-2.6.jar wordmedian /files/file327.txt /wordmedian_out
[acadgild.mmisra home]$
18/07/03 15:19:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where ap                              plicable
18/07/03 15:19:25 INFO client.RMProxy: Connecting to ResourceManager at
localhost/127.0.0.1:8032
18/07/03 15:19:26 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing
not performed. Implement the Tool interface and ex                           ecute
your application with ToolRunner to remedy this.
18/07/03 15:19:26 INFO input.FileInputFormat: Total input paths to process : 1
18/07/03 15:19:26 INFO mapreduce.JobSubmitter: number of splits:3
18/07/03 15:19:26 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1530595889847_0002
18/07/03 15:19:26 INFO impl.YarnClientImpl: Submitted application
application_1530595889847_0002
18/07/03 15:19:26 INFO mapreduce.Job: The url to track the job:
http://localhost:8088/proxy/application_1530595889847_0002/
18/07/03 15:19:26 INFO mapreduce.Job: Running job: job_1530595889847_0002
18/07/03 15:19:34 INFO mapreduce.Job: Job job_1530595889847_0002 running in uber mode :
false
18/07/03 15:19:34 INFO mapreduce.Job:  map 0% reduce 0%
18/07/03 15:19:49 INFO mapreduce.Job:  map 3% reduce 0%
18/07/03 15:19:52 INFO mapreduce.Job:  map 11% reduce 0%
18/07/03 15:19:53 INFO mapreduce.Job:  map 14% reduce 0%
18/07/03 15:19:56 INFO mapreduce.Job:  map 15% reduce 0%
18/07/03 15:19:58 INFO mapreduce.Job:  map 20% reduce 0%
18/07/03 15:19:59 INFO mapreduce.Job:  map 22% reduce 0%
18/07/03 15:20:01 INFO mapreduce.Job:  map 26% reduce 0%
18/07/03 15:20:02 INFO mapreduce.Job:  map 28% reduce 0%
18/07/03 15:20:04 INFO mapreduce.Job:  map 34% reduce 0%
18/07/03 15:20:05 INFO mapreduce.Job:  map 35% reduce 0%
18/07/03 15:20:07 INFO mapreduce.Job:  map 41% reduce 0%
18/07/03 15:20:08 INFO mapreduce.Job:  map 43% reduce 0%
18/07/03 15:20:10 INFO mapreduce.Job:  map 47% reduce 0%
18/07/03 15:20:11 INFO mapreduce.Job:  map 60% reduce 0%
18/07/03 15:20:13 INFO mapreduce.Job:  map 63% reduce 0%
18/07/03 15:20:14 INFO mapreduce.Job:  map 65% reduce 0%
18/07/03 15:20:16 INFO mapreduce.Job:  map 66% reduce 0%
18/07/03 15:20:17 INFO mapreduce.Job:  map 67% reduce 0%
18/07/03 15:20:19 INFO mapreduce.Job:  map 69% reduce 0%
18/07/03 15:20:20 INFO mapreduce.Job:  map 71% reduce 0%
18/07/03 15:20:23 INFO mapreduce.Job:  map 72% reduce 0%
18/07/03 15:20:25 INFO mapreduce.Job:  map 74% reduce 0%
18/07/03 15:20:26 INFO mapreduce.Job:  map 76% reduce 0%
18/07/03 15:20:27 INFO mapreduce.Job:  map 76% reduce 11%
18/07/03 15:20:28 INFO mapreduce.Job:  map 87% reduce 11%
18/07/03 15:20:29 INFO mapreduce.Job:  map 89% reduce 11%
18/07/03 15:20:30 INFO mapreduce.Job:  map 100% reduce 22%
18/07/03 15:20:31 INFO mapreduce.Job:  map 100% reduce 100%
18/07/03 15:20:31 INFO mapreduce.Job: Job job_1530595889847_0002 completed successfully
18/07/03 15:20:32 INFO mapreduce.Job: Counters: 50
        File System Counters
                FILE: Number of bytes read=1506
                FILE: Number of bytes written=431885
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
```

```
                HDFS: Number of bytes read=327164216
                HDFS: Number of bytes written=61
                HDFS: Number of read operations=12
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Killed map tasks=1
                Launched map tasks=4
                Launched reduce tasks=1
                Data-local map tasks=4
                Total time spent by all maps in occupied slots (ms)=159236
                Total time spent by all reduces in occupied slots (ms)=17582
                Total time spent by all map tasks (ms)=159236
                Total time spent by all reduce tasks (ms)=17582
                Total vcore-milliseconds taken by all map tasks=159236
                Total vcore-milliseconds taken by all reduce tasks=17582
                Total megabyte-milliseconds taken by all map tasks=163057664
                Total megabyte-milliseconds taken by all reduce tasks=18003968
        Map-Reduce Framework
                Map input records=8388608
                Map output records=67108864
                Map output bytes=536870912
                Map output materialized bytes=198
                Input split bytes=312
                Combine input records=67108984
                Combine output records=138
                Reduce input groups=6
                Reduce shuffle bytes=198
                Reduce input records=18
                Reduce output records=6
                Spilled Records=156
                Shuffled Maps =3
                Failed Shuffles=0
                Merged Map outputs=3
                GC time elapsed (ms)=1052
                CPU time spent (ms)=68690
                Physical memory (bytes) snapshot=952061952
                Virtual memory (bytes) snapshot=8328830976
                Total committed heap usage (bytes)=728760320
        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=327163904
        File Output Format Counters
                Bytes Written=61
The median is: 3


#check the reducer output in the /wordmedian_out directory
 [acadgild.mmisra home]$
[acadgild.mmisra home]$ hadoop fs -ls /
18/07/03 15:21:11 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 9 items
drwxr-xr-x   - acadgild supergroup          0 2018-06-24 10:51 /files
drwxr-xr-x   - acadgild supergroup          0 2018-06-24 10:50 /mohit
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 09:03 /out1
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 10:01 /sqoopout
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 11:21 /sqoopoutbyid
```

```
drwx------   - acadgild supergroup          0 2018-06-24 11:27 /tmp
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 14:12 /user
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 15:17 /wordmean_out
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 15:20 /wordmedian_out
[acadgild.mmisra home]$ hadoop fs -ls /wordmedian_out
18/07/03 15:21:20 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2018-07-03 15:20 /wordmedian_out/_SUCCESS
-rw-r--r--   1 acadgild supergroup         61 2018-07-03 15:20
/wordmedian_out/part-r-00000


[acadgild.mmisra home]$ hadoop fs -cat /wordmedian_out/part-r-00000
18/07/03 15:21:39 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where ap                                plicable
1       8388608
2       8388608
3       25165824
4       8388608
7       8388608
8       8388608
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra home]$
[acadgild.mmisra home]$
[acadgild.mmisra home]$ hadoop jar
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa
mples-2.6.jar
An example program must be given as the first argument.
Valid program names are:
  aggregatewordcount: An Aggregate based map/reduce program that counts the words in the
input files.
  aggregatewordhist: An Aggregate based map/reduce program that computes the histogram of
the words in the input files.
  bbp: A map/reduce program that uses Bailey-Borwein-Plouffe to compute exact digits of
Pi.
  dbcount: An example job that count the pageview counts from a database.
  distbbp: A map/reduce program that uses a BBP-type formula to compute exact bits of Pi.
  grep: A map/reduce program that counts the matches of a regex in the input.
  join: A job that effects a join over sorted, equally partitioned datasets
  multifilewc: A job that counts words from several files.
  pentomino: A map/reduce tile laying program to find solutions to pentomino problems.
  pi: A map/reduce program that estimates Pi using a quasi-Monte Carlo method.
  randomtextwriter: A map/reduce program that writes 10GB of random textual data per node.
  randomwriter: A map/reduce program that writes 10GB of random data per node.
  secondarysort: An example defining a secondary sort to the reduce.
  sort: A map/reduce program that sorts the data written by the random writer.
  sudoku: A sudoku solver.
  teragen: Generate data for the terasort
  terasort: Run the terasort
  teravalidate: Checking results of terasort
  wordcount: A map/reduce program that counts the words in the input files.
  wordmean: A map/reduce program that counts the average length of the words in the input
files.
  wordmedian: A map/reduce program that counts the median length of the words in the input
files.
  wordstandarddeviation: A map/reduce program that counts the standard deviation of the
length of the words in the input files.
```

 # Run the wordstandardeviation option which is a map/reduce program that counts the standard deviation of the length of the words in the input files We run the program on the same file327.txt and the output is stored on HDFS under /wordsd_out directory

[acadgild.mmisra home]$ **hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exa mples-2.6.jar wordstandarddeviation /files/files327.txt /wordsd_out**

18/07/03 15:26:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/03 15:26:47 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/03 15:26:48 INFO input.FileInputFormat: Total input paths to process : 1
18/07/03 15:26:48 INFO mapreduce.JobSubmitter: number of splits:3
18/07/03 15:26:49 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530595889847_0004
18/07/03 15:26:49 INFO impl.YarnClientImpl: Submitted application application_1530595889847_0004
18/07/03 15:26:49 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1530595889847_0004/
18/07/03 15:26:49 INFO mapreduce.Job: Running job: job_1530595889847_0004
18/07/03 15:26:56 INFO mapreduce.Job: Job job_1530595889847_0004 running in uber mode : false
18/07/03 15:26:56 INFO mapreduce.Job: map 0% reduce 0%
18/07/03 15:27:11 INFO mapreduce.Job: map 6% reduce 0%
18/07/03 15:27:14 INFO mapreduce.Job: map 9% reduce 0%
18/07/03 15:27:17 INFO mapreduce.Job: map 12% reduce 0%
18/07/03 15:27:21 INFO mapreduce.Job: map 16% reduce 0%
18/07/03 15:27:24 INFO mapreduce.Job: map 20% reduce 0%
18/07/03 15:27:27 INFO mapreduce.Job: map 23% reduce 0%
18/07/03 15:27:30 INFO mapreduce.Job: map 27% reduce 0%
18/07/03 15:27:33 INFO mapreduce.Job: map 31% reduce 0%
18/07/03 15:27:36 INFO mapreduce.Job: map 34% reduce 0%
18/07/03 15:27:37 INFO mapreduce.Job: map 35% reduce 0%
18/07/03 15:27:39 INFO mapreduce.Job: map 37% reduce 0%
18/07/03 15:27:40 INFO mapreduce.Job: map 38% reduce 0%
18/07/03 15:27:42 INFO mapreduce.Job: map 40% reduce 0%
18/07/03 15:27:43 INFO mapreduce.Job: map 41% reduce 0%
18/07/03 15:27:45 INFO mapreduce.Job: map 43% reduce 0%
18/07/03 15:27:46 INFO mapreduce.Job: map 55% reduce 0%
18/07/03 15:27:48 INFO mapreduce.Job: map 56% reduce 0%
18/07/03 15:27:49 INFO mapreduce.Job: map 57% reduce 0%
18/07/03 15:27:52 INFO mapreduce.Job: map 58% reduce 0%
18/07/03 15:27:57 INFO mapreduce.Job: map 59% reduce 0%
18/07/03 15:28:01 INFO mapreduce.Job: map 61% reduce 0%
18/07/03 15:28:04 INFO mapreduce.Job: map 62% reduce 0%
18/07/03 15:28:06 INFO mapreduce.Job: map 62% reduce 11%
18/07/03 15:28:07 INFO mapreduce.Job: map 63% reduce 11%
18/07/03 15:28:10 INFO mapreduce.Job: map 64% reduce 11%
18/07/03 15:28:13 INFO mapreduce.Job: map 65% reduce 11%
18/07/03 15:28:16 INFO mapreduce.Job: map 67% reduce 11%
18/07/03 15:28:19 INFO mapreduce.Job: map 68% reduce 11%
18/07/03 15:28:22 INFO mapreduce.Job: map 69% reduce 11%
18/07/03 15:28:25 INFO mapreduce.Job: map 70% reduce 11%
18/07/03 15:28:28 INFO mapreduce.Job: map 71% reduce 11%
18/07/03 15:28:31 INFO mapreduce.Job: map 73% reduce 11%
18/07/03 15:28:34 INFO mapreduce.Job: map 74% reduce 11%
18/07/03 15:28:37 INFO mapreduce.Job: map 76% reduce 11%
18/07/03 15:28:40 INFO mapreduce.Job: map 77% reduce 11%
18/07/03 15:28:43 INFO mapreduce.Job: map 89% reduce 11%
18/07/03 15:28:45 INFO mapreduce.Job: map 100% reduce 100%

```
18/07/03 15:28:46 INFO mapreduce.Job: Job job_1530595889847_0004 completed successfully
18/07/03 15:28:46 INFO mapreduce.Job: Counters: 50
        File System Counters
                FILE: Number of bytes read=7104
                FILE: Number of bytes written=438285
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=327164216
                HDFS: Number of bytes written=50
                HDFS: Number of read operations=12
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Killed map tasks=2
                Launched map tasks=5
                Launched reduce tasks=1
                Data-local map tasks=5
                Total time spent by all maps in occupied slots (ms)=356040
                Total time spent by all reduces in occupied slots (ms)=56905
                Total time spent by all map tasks (ms)=356040
                Total time spent by all reduce tasks (ms)=56905
                Total vcore-milliseconds taken by all map tasks=356040
                Total vcore-milliseconds taken by all reduce tasks=56905
                Total megabyte-milliseconds taken by all map tasks=364584960
                Total megabyte-milliseconds taken by all reduce tasks=58270720
        Map-Reduce Framework
                Map input records=8388608
                Map output records=201326592
                Map output bytes=2952790016
                Map output materialized bytes=168
                Input split bytes=312
                Combine input records=201326820
                Combine output records=237
                Reduce input groups=3
                Reduce shuffle bytes=168
                Reduce input records=9
                Reduce output records=3
                Spilled Records=405
                Shuffled Maps =3
                Failed Shuffles=0
                Merged Map outputs=3
                GC time elapsed (ms)=1051
                CPU time spent (ms)=130380
                Physical memory (bytes) snapshot=942788608
                Virtual memory (bytes) snapshot=8331960320
                Total committed heap usage (bytes)=684195840
        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=327163904
        File Output Format Counters
                Bytes Written=50
```
The standard deviation is: 2.2603926650031405
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra home]$
[acadgild.mmisra home]$

```
# check the redeucer output in /wordsd_out dir
[acadgild.mmisra home]$ hadoop fs -ls /
18/07/03 15:29:29 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 10 items
drwxr-xr-x   - acadgild supergroup          0 2018-06-24 10:51 /files
drwxr-xr-x   - acadgild supergroup          0 2018-06-24 10:50 /mohit
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 09:03 /out1
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 10:01 /sqoopout
drwxr-xr-x   - acadgild supergroup          0 2018-07-01 11:21 /sqoopoutbyid
drwx------   - acadgild supergroup          0 2018-06-24 11:27 /tmp
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 14:12 /user
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 15:17 /wordmean_out
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 15:20 /wordmedian_out
drwxr-xr-x   - acadgild supergroup          0 2018-07-03 15:28 /wordsd_out
[acadgild.mmisra home]$ hadoop fs -ls /wordsd_out
18/07/03 15:29:43 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2018-07-03 15:28 /wordsd_out/_SUCCESS
-rw-r--r--   1 acadgild supergroup         50 2018-07-03 15:28 /wordsd_out/part-r-00000
[acadgild.mmisra home]$ hadoop fs -cat /wordsd_out/part-r-00000
18/07/03 15:29:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
count   67108864
length  260046848
square  1350565888
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra home]$
[acadgild.mmisra home]$
[acadgild.mmisra home]$
[acadgild.mmisra home]$
```