

Assignment 3:Map-reduce

1.WordMedian :

Comand:

hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian /user/acadgild/input/abc.txt
/user/acadgild/output/wordmedian/

```
File Edit View Search Terminal Help
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian /user/acadgild/input/abc.txt /user/acadgild/output/wordmedian/
```

create _Sucess and part-r-00000 file in /user/acadgild/output/wordmedian

```
[acadgild@localhost mapreduce]$ hdfs dfs -ls /user/acadgild/output/wordmedian/
18/09/01 12:57:12 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 2018-09-01 12:56 /user/acadgild/output/wordmedian/ SUCCESS
-rw-r--r-- 1 acadgild supergroup 20 2018-09-01 12:56 /user/acadgild/output/wordmedian/part-r-00000
```

explanation:

A mapreduce program that counts the median length of the words in the input files.

Output:

```
[acadgild@localhost mapreduce]$ hdfs dfs -cat /user/acadgild/output/wordmedian/part-r-00000
18/09/01 13:00:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
1      1
2      2
4      6
5      5
6      4
```

wordmedian for “acadgild” in sample.txt file .

```
The median is: 8
[acadgild@localhost mapreduce]$ hdfs dfs -cat /user/acadgild/output/wordmedian1/part-r-00000
18/09/01 13:05:13 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
8      1
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$
```

2. WordMean:

Command :

hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean /user/acadgild/input/sample.txt
/user/acadgild/output/wordmean/

Screenshot:

```
[acadgild@localhost mapreduce]$
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean /user/acadgild/input/sample.txt /user/acadgild/output/wordmean/
18/09/01 13:09:09 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/09/01 13:09:11 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/09/01 13:09:12 INFO input.FileInputFormat: Total input paths to process : 1
18/09/01 13:09:13 INFO mapreduce.JobSubmitter: number of splits:1
18/09/01 13:09:13 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1535781261898_0003
18/09/01 13:09:13 INFO impl.YarnClientImpl: Submitted application application_1535781261898_0003
18/09/01 13:09:13 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1535781261898_0003/
18/09/01 13:09:13 INFO mapreduce.Job: Running job: job_1535781261898_0003
18/09/01 13:09:25 INFO mapreduce.Job: Job job_1535781261898_0003 running in uber mode : false
18/09/01 13:09:25 INFO mapreduce.Job: map 0% reduce 0%
18/09/01 13:09:33 INFO mapreduce.Job: map 100% reduce 0%
18/09/01 13:09:42 INFO mapreduce.Job: map 100% reduce 100%
18/09/01 13:09:42 INFO mapreduce.Job: Job job_1535781261898_0003 completed successfully
18/09/01 13:09:42 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=39
    FILE: Number of bytes written=215469
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=126
    HDFS: Number of bytes written=17
    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
```

Explanation of Command:

wordmean: A mapreduce program that counts the average length of the words in the input files.

Output:

```
[acadgild@localhost mapreduce]$ hdfs dfs -cat /user/acadgild/output/wordmean/part-r-000000
18/09/01 13:15:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
count 1
length 8
You have new mail in /var/spool/mail/acadgild
```

3. WordStandardDeviation

Command:

`hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /user/acadgild/input/sample.txt /user/acadgild/output/wordstandarddeviation/`

```
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /user/acadgild/input/sample.txt /user/acadgild/output/wordstandarddeviation/
18/09/01 13:19:07 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/09/01 13:19:08 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/09/01 13:19:10 INFO input.FileInputFormat: Total input paths to process : 1
18/09/01 13:19:10 INFO mapreduce.JobSubmitter: number of splits:1
18/09/01 13:19:10 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1535781261898_0004
18/09/01 13:19:11 INFO impl.YarnClientImpl: Submitted application application_1535781261898_0004
18/09/01 13:19:11 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1535781261898_0004/
18/09/01 13:19:11 INFO mapreduce.Job: Running job: job_1535781261898_0004
18/09/01 13:19:21 INFO mapreduce.Job: Job job_1535781261898_0004 running in uber mode : false
18/09/01 13:19:21 INFO mapreduce.Job: map 0% reduce 0%
18/09/01 13:19:30 INFO mapreduce.Job: map 100% reduce 0%
18/09/01 13:19:39 INFO mapreduce.Job: map 100% reduce 100%
18/09/01 13:19:39 INFO mapreduce.Job: Job job_1535781261898_0004 completed successfully
18/09/01 13:19:39 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=56
    FILE: Number of bytes written=215689
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=126
    HDFS: Number of bytes written=27
    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)=5706
    Total time spent by all reduces in occupied slots (ms)=6150
    Total time spent by all map tasks (ms)=5706
    Total time spent by all reduce tasks (ms)=6150
```

Explanation of Command:

wordstandarddeviation: A mapreduce program that counts the standard deviation of the length of the words in the input files.

Output:

```
[acadgild@localhost mapreduce]$ hdfs dfs -cat /user/acadgild/output/wordstandarddeviation/part-r-00000
18/09/01 13:21:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
count      1
length     8
square    64
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$
```