Session 3: YARN

Assignment 3

Execute WordMedian , WordMean , WordStandardDeviation programs using

hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM

- To execute the given commands, we are using the file word-count.txt, whose contents are as shown below:

```
[acadgild@localhost bin]$ hadoop fs -cat /user/acadgild/hadoop/word-count.txt
18/12/03 23:58:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
This is for testing if we can enter text in the file using appendToFile command in hadoop. Once done typing, we press Ctrl+D
to exit the editor.[acadgild@localhost bin]$
```

1. WordMedian command : A map/reduce program that counts the median length of the words in the input files.

```
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost bin]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examp
les-2.6.5.jar wordmedian /user/acadgild/hadoop/word-count.txt /user/wordmedianoutput
18/12/04 00:07:00 WARN util NativeCodeLoader: Unable to load native hadoop library for your platform    using builtin java cl
```

The result of execution of the command is as shown:

```
        BAD_ID=0
        CONNECTION=0
        IO_ERROR=0
        WRONG_LENGTH=0
        WRONG_MAP=0
        WRONG_REDUCE=0
    File Input Format Counters
        Bytes Read=144
    File Output Format Counters
        Bytes Written=29
The median is: 4
[acadgild@localhost bin]$
```

Also, the output directory wordmedianoutput gets created in the path mentioned while running, i.e /user/wordmedianoutput, which we can see in the browser as shown below:

Hadoop, 2016.

2. **WordMean :** A map/reduce program that counts the average length of the words in the input files.

```
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost bin]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examp
les-2.6.5.jar wordmean /user/acadgild/hadoop/word-count.txt /user/wordmeanoutput
```

The output of the command is as follows:


```
            File Input Format Counters
                    Bytes Read=144
            File Output Format Counters
                    Bytes Written=20
The mean is: 4.37037037037037
```

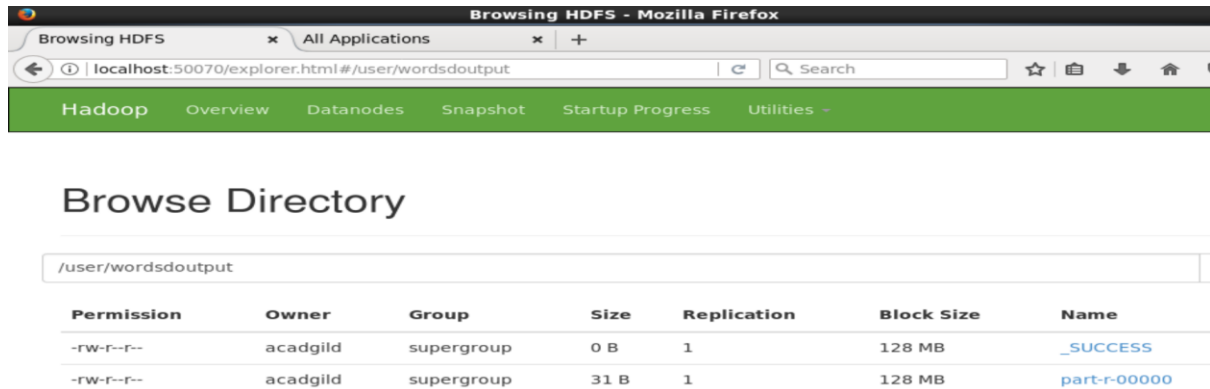Also, the directory created in HDFS is as shown below:

3. **WordStandardDeviation command :** 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
[acadgild@localhost bin]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examp
les-2.6.5.jar wordstandarddeviation /user/acadgild/hadoop/word-count.txt /user/wordsdoutput
18/12/04 00:23:08 WARN util NativeCodeLoader: Unable to load native-hadoop library for your platform    using builtin-java cl
```

The output of the command is as shown below:

```
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=144
        File Output Format Counters
                Bytes Written=31
The standard deviation is: 2.3117757828632937
```

The output directory created and browsed as shown below:

Browsing HDFS × All Applications × +

localhost:50070/explorer.html#/user/wordsdoutput

Hadoop   Overview   Datanodes   Snapshot   Startup Progress   Utilities

Browse Directory

/user/wordsdoutput

Permission	Owner	Group	Size	Replication	Block Size	Name
-rw-r--r--	acadgild	supergroup	0 B	1	128 MB	_SUCCESS
-rw-r--r--	acadgild	supergroup	31 B	1	128 MB	part-r-00000