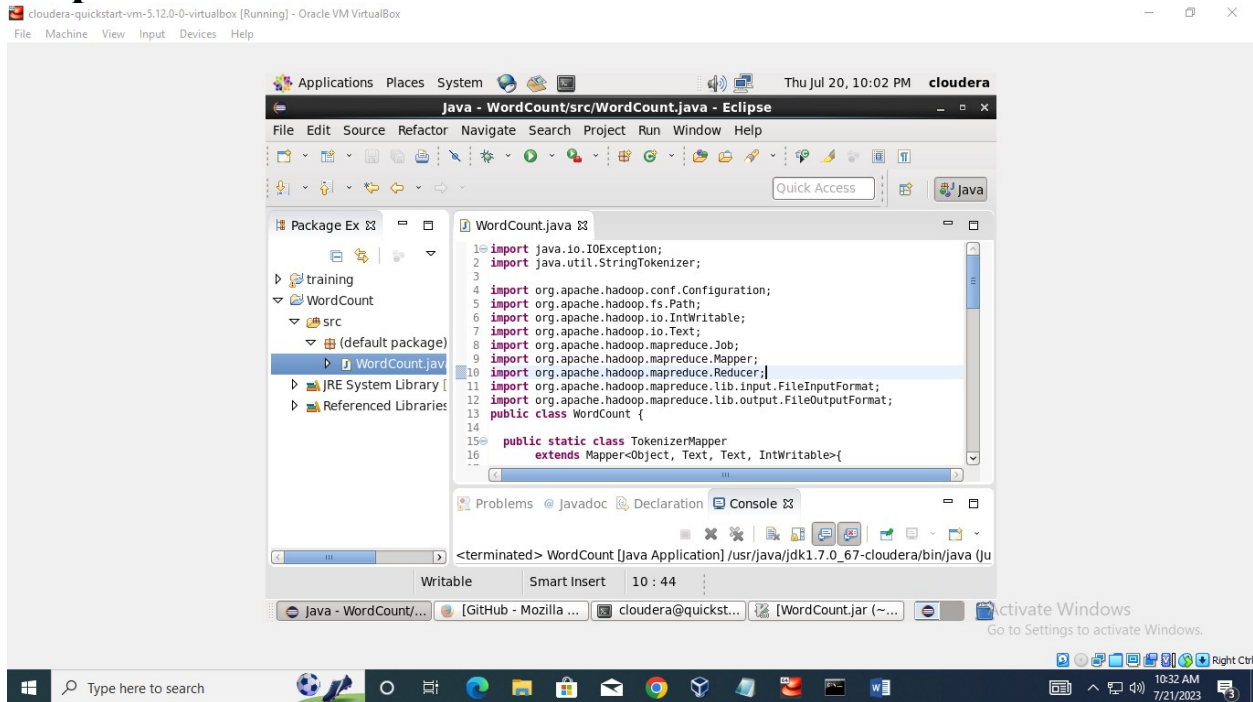
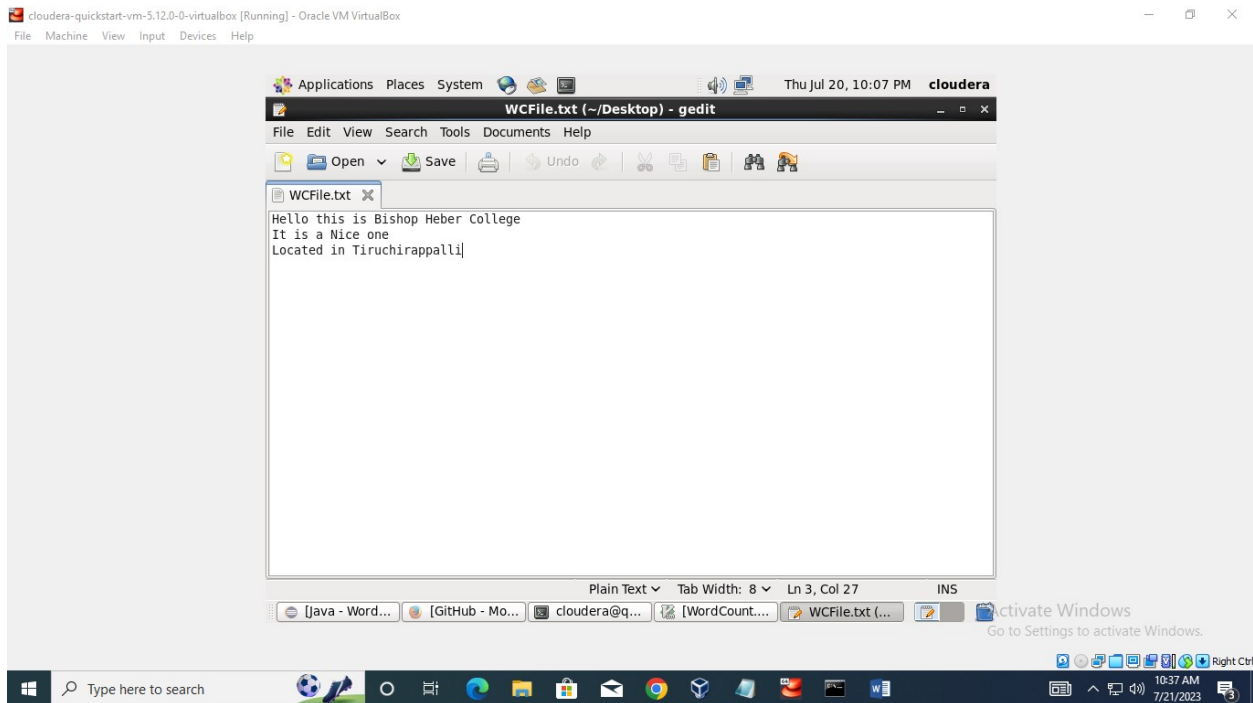


PASUPATHEESWARAN

Step 1:





```
[cloudera@quickstart Desktop]$ hdfsdfs -put home/cloudera/Desktop/WCFFile.txt /input/
```

```
[cloudera@quickstart Desktop]$ hdfsdfs -ls /input/
```

Found 1 items

```
-rw-r--r--  1 clouderasupergroup    79 2023-07-20 21:38 /input/WCFFile.txt
```

Step 4:

```
[cloudera@quickstart Desktop]$ hadoop jar /home/cloudera/Desktop/WordCount.jar WordCount /input/WCFFile.txt /out00
```

```
23/07/20 21:53:00 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
```

```
23/07/20 21:53:00 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
```

```
23/07/20 21:53:00 WARN hdfs.DFSCClient: Caught exception
```

```
java.lang.InterruptedExcep
```

```
atjava.lang.Object.wait(Native Method)
```

```
atjava.lang.Thread.join(Thread.java:1281)
```

```
    atjava.lang.Thread.join(Thread.java:1355)

    at
org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.closeResponder(DFSOutputStream.java:952)

    at
org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.endBlock(DFSOutputStream.java:690)
)

    at
org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.run(DFSOutputStream.java:879)
```

23/07/20 21:53:00 INFO input.FileInputFormat: Total input paths to process : 1

23/07/20 21:53:00 INFO mapreduce.JobSubmitter: number of splits:1

23/07/20 21:53:01 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1689910151412_0001

23/07/20 21:53:01 INFO impl.YarnClientImpl: Submitted application
application_1689910151412_0001

23/07/20 21:53:01 INFO mapreduce.Job: The url to track the job:
http://quickstart.cloudera:8088/proxy/application_1689910151412_0001/

23/07/20 21:53:01 INFO mapreduce.Job: Running job: job_1689910151412_0001

23/07/20 21:53:11 INFO mapreduce.Job: Job job_1689910151412_0001 running in ubermode :
false

23/07/20 21:53:11 INFO mapreduce.Job: map 0% reduce 0%

23/07/20 21:53:18 INFO mapreduce.Job: map 100% reduce 0%

23/07/20 21:53:23 INFO mapreduce.Job: map 100% reduce 100%

23/07/20 21:53:23 INFO mapreduce.Job: Job job_1689910151412_0001 completed successfully

23/07/20 21:53:23 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=160

FILE: Number of bytes written=250291

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=192

HDFS: Number of bytes written=102

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)=4052

Total time spent by all reduces in occupied slots (ms)=3071

Total time spent by all map tasks (ms)=4052

Total time spent by all reduce tasks (ms)=3071

Total vcore-milliseconds taken by all map tasks=4052

Total vcore-milliseconds taken by all reduce tasks=3071

Total megabyte-milliseconds taken by all map tasks=4149248

Total megabyte-milliseconds taken by all reduce tasks=3144704

Map-Reduce Framework

Map input records=3

Map output records=14

Map output bytes=135

Map output materialized bytes=160

Input split bytes=113

Combine input records=14

Combine output records=13

Reduce input groups=13

Reduce shuffle bytes=160

Reduce input records=13

Reduce output records=13

Spilled Records=26

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=96

CPU time spent (ms)=750

Physical memory (bytes) snapshot=341028864

Virtual memory (bytes) snapshot=3015155712

Total committed heap usage (bytes)=226365440

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=79

File Output Format Counters

Bytes Written=102

Step 5:

```
[cloudera@quickstart Desktop]$ hdfsdfs -ls /out00
```

Found 2 items

```
-rw-r--r--  1 clouderasupergroup      0 2023-07-20 21:53 /out00/_SUCCESS
```

```
-rw-r--r--  1 clouderasupergroup    102 2023-07-20 21:53 /out00/part-r-000000
```

```
[cloudera@quickstart Desktop]$ hdfsdfs -ls /out00/part-r-000000
```

```
-rw-r--r--  1 clouderasupergroup    102 2023-07-20 21:53 /out00/part-r-000000
```

```
[cloudera@quickstart Desktop]$ hdfsdfs /out00/part-r-000000
```

/out00/part-r-000000: Unknown command

```
[cloudera@quickstart Desktop]$ hdfsdfs -cat /out00/part-r-000000
```

Bishop 1

College 1

Heber 1

Hello 1

It 1

Located 1

Nice 1

Tiruchirappalli1

a 1

in 1

is 2

one 1

this 1