

# Report for Lab Assignment 4

## Question 1

### Description:

Hadoop MapReduce Algorithm – To implement a MapReduce algorithm to find the Facebook Mutual Friends and to run this Map Reduce job on Apache Hadoop.

### Snapshot:

Running Map Reduce job in cloudera.

The screenshot shows a Cloudera Quickstart VM running in a VMware Workstation 12 Player. The desktop environment includes icons for Computer, Eclipse, Big Data, Trash, cloudera's Home, Mutual, cm, Launch Cloudera Express, Launch Cloudera Enterprise (trial), and Configure Mirrors. A terminal window is open, displaying the following commands and output:

```
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
WRONG MAP=0
WRONG REDUCE=0
File Input Format Counters
Bytes Read=49
File Output Format Counters
Bytes Written=57
[cloudera@quickstart Desktop]$ hadoop fs -cat output/*
cat: 'output/*': No such file or directory
[cloudera@quickstart Desktop]$ hadoop fs -cat Output/*
cat: Output/*: No such file or directory
[cloudera@quickstart Desktop]$ hadoop fs -cat Output.txt/*
AB CD
AC BD
AD BC
BC ADE
BD ACE
BE CD
CD ABE
CE BD
DE BC
[cloudera@quickstart Desktop]$ hadoop fs -put Mutual.txt Input.txt
put: Input.txt: File exists
[cloudera@quickstart Desktop]$ hadoop fs -put Mutual.txt Input1.txt
[cloudera@quickstart Desktop]$
```

In the background, a presentation slide titled "Input File" is visible, containing the following text:

**Input File**

- Create a local file in terminal  
*vi input*
- Push the file to Hadoop  
*hadoop fs -put localinput hadoopinput*
- *hadoop fs -ls*  
*[group1@kc-sce-cloudera-01 ~]\$ hadoo*  
*Found 5 items*

<i>drwx-----</i>	<i>- group1 supergroup</i>	<i>0</i>
<i>drwx-----</i>	<i>- group1 supergroup</i>	<i>0</i>
<i>-rw-r--r--</i>	<i>3 group1 supergroup</i>	<i>15</i>
<i>drwxr-xr-x</i>	<i>- group1 supergroup</i>	<i>0</i>
<i>drwxr-xr-x</i>	<i>- group1 supergroup</i>	<i>0</i>

The bottom of the screen shows the Windows taskbar with the date 2/17/2016 and time 8:12 PM.

Google Account | CS5542 Big Data | Dropbox - CS5 | Dropbox - CS5 | Dropbox - CS5 | Maven - Down | Welcome, Priy | Mail - Nidad | Apache Hado | Step 11: Start | How-to: Use | Other bookmarks

CS5542-Tutorial4-Hadoop Map\_Reduce.pptx

# Input File

- Create a local input file
- Push the file to HDFS
- `hadoop fs -put Input.txt /mutual`
- `hadoop fs -ls /mutual`

Found 5 items

```
drwx----- -g cloudera cloudera 49 2016-02-17 15:22 Input.txt
drwx----- -g cloudera cloudera 49 2016-02-17 18:12 Input1.txt
-rw-r--r-- 3 cloudera cloudera 0 2016-02-17 15:26 Output.txt
drwxr-xr-x -g cloudera cloudera 49 2016-02-17 15:26 Output.txt
```

cloudera@quickstart:~/Desktop

```
File Edit View Search Terminal Help
Bytes Written=57
[cloudera@quickstart Desktop]$ hadoop fs -cat output/*
cat: 'output/*': No such file or directory
[cloudera@quickstart Desktop]$ hadoop fs -cat Output/*
cat: 'Output/*': No such file or directory
[cloudera@quickstart Desktop]$ hadoop fs -cat Output.txt/*
AB
CD
AC
BD
AD
BC
ADE
ACE
BE
CE
ABE
BC
DE
[cloudera@quickstart Desktop]$ hadoop fs -put Mutual.txt Input.txt
put: 'Input.txt': File exists
[cloudera@quickstart Desktop]$ hadoop fs -put Mutual.txt Input1.txt
[cloudera@quickstart Desktop]$ hadoop fs -ls
Found 3 items
-rw-r--r-- 1 cloudera cloudera 49 2016-02-17 15:22 Input.txt
-rw-r--r-- 1 cloudera cloudera 49 2016-02-17 18:12 Input1.txt
drwxr-xr-x - cloudera cloudera 0 2016-02-17 15:26 Output.txt
[cloudera@quickstart Desktop]$
```

Big Data | 0 items, Free space: 43.9 GB

Connecting to cloudera | Downloads | Big Data | Computer | / | Java - trainin... | cloudera@qui...

8:12 PM 2/17/2016

Google Account | CS5542 Big Data | Dropbox - CS5 | Dropbox - CS5 | Dropbox - CS5 | Maven - Down | Welcome, Priy | Mail - Nidad | Apache Hado | Step 11: Start | How-to: Use | Other bookmarks

CS5542-Tutorial4-Hadoop Map\_Reduce.pptx

# Run Hadoop WordCount

```
hadoop jar jarfile MainClass Input output
```

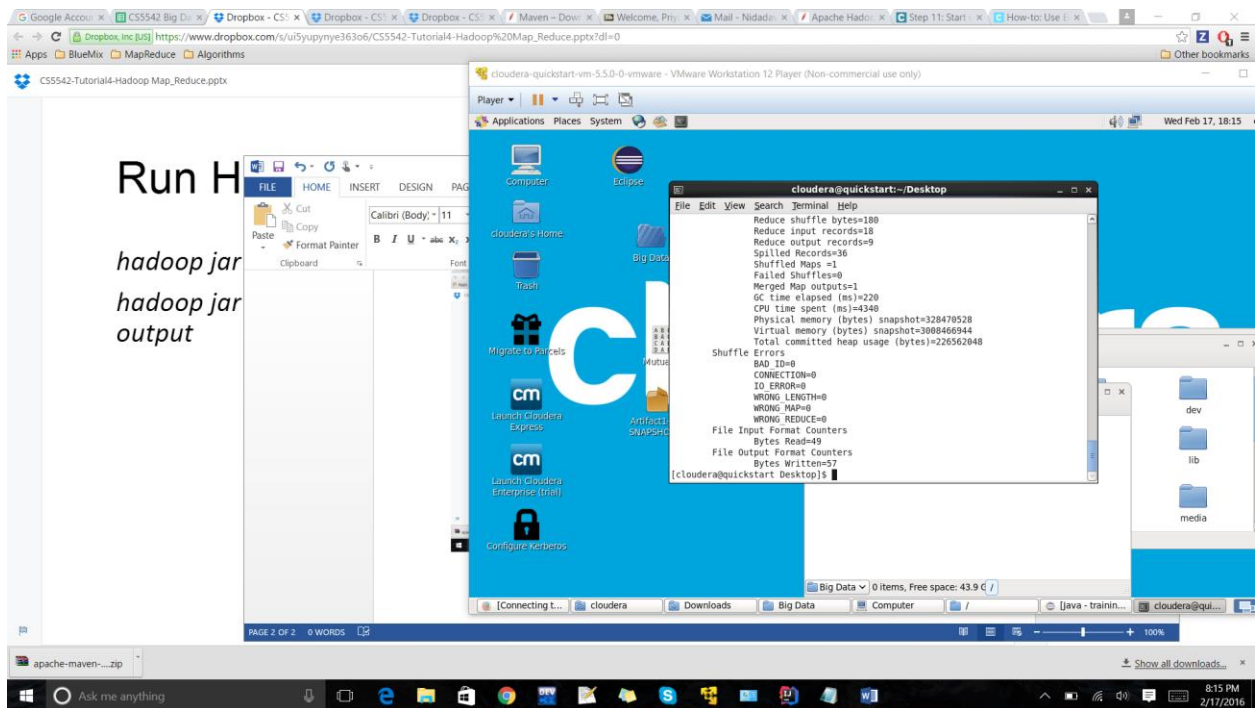
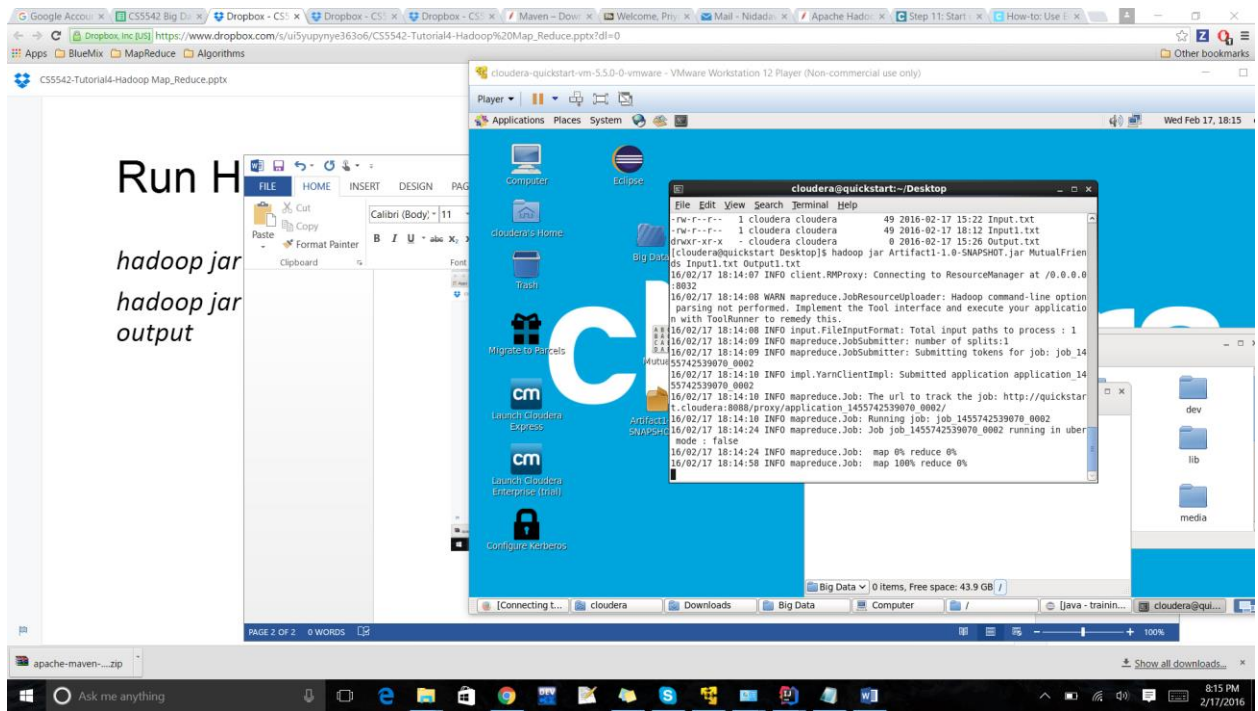
cloudera@quickstart:~/Desktop

```
File Edit View Search Terminal Help
[cloudera@quickstart Desktop]$ hadoop fs -put Mutual.txt Input.txt
put: 'Input.txt': File exists
[cloudera@quickstart Desktop]$ hadoop fs -put Mutual.txt Input1.txt
[cloudera@quickstart Desktop]$ hadoop fs -ls
Found 3 items
-rw-r--r-- 1 cloudera cloudera 49 2016-02-17 15:22 Input.txt
-rw-r--r-- 1 cloudera cloudera 49 2016-02-17 18:12 Input1.txt
drwxr-xr-x - cloudera cloudera 0 2016-02-17 15:26 Output.txt
[cloudera@quickstart Desktop]$ hadoop jar Artifact1-1.0-SNAPSHOT.jar MutualFriends Input1.txt Output1.txt
16/02/17 18:14:07 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
16/02/17 18:14:08 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
16/02/17 18:14:08 INFO input.FileInputFormat: Total input paths to process : 1
16/02/17 18:14:09 INFO mapreduce.JobSubmitter: number of splits:1
16/02/17 18:14:09 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1455742539870_0002
```

Big Data | 0 items, Free space: 43.9 GB

Connecting to cloudera | Downloads | Big Data | Computer | / | Java - trainin... | cloudera@qui...

8:14 PM 2/17/2016



Google Account CS5542 Big Data Dropbox - CS: X Dropbox - CS: X Dropbox - CS: X Maven - Down: X Welcome, Pri: X Mail - Nidada: X Apache Hado: X Step 11: Start: X How-to: Use E: X

CS5542-Tutorial4-Hadoop Map\_Reduce.pptx

# Check output

- `hadoop fs -cat output/*`

[group1@kc-sce-cloudera-01 ~]

```
Hi 1
are 1
how 1
you 1
```

cloudera-quickstart-vm-5.5.0-0-vmware - VMware Workstation 12 Player (Non-commercial use only)

Applications Places System Wed Feb 17, 18:16

cloudera@quickstart:~/Desktop

```
File Edit View Search Terminal Help
Virtual memory (bytes) snapshot=3008406944
Total committed heap usage (bytes)=226562048
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=49
File Output Format Counters
  Bytes Written=57
[cloudera@quickstart Desktop]$ hadoop fs -cat Output1.txt/*
AB
CD
AC
BD
AD
BC
BC
ADE
BD
ACE
BE
CD
ADE
CE
BD
DE
BC
```

Big Data 0 items, Free space: 43.9 G /

ask me anything 8:16 PM 2/17/2016

## Question 2

### Description:

Android Application – To implement a smart phone application which uses speech services.

I had implemented the speech to text service. Using this, I can make my Robot to interact with me.

### Snapshot:

