

Lab 7 Word Count using Map-Reduce

```
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [bmscece-HP-Elite-Tower-800-G9-Desktop-PC]
Starting resourcemanager
Starting nodemanagers
```

Hadoop services are started using start-all.sh, launching daemons like NameNode, DataNode, and ResourceManager.

```
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~$ jps
7042 DataNode
7639 ResourceManager
8248 Jps
6904 NameNode
7305 SecondaryNameNode
7788 NodeManager
4975 org.eclipse.equinox.launcher_1.6.1000.v20250227-1734.jar
```

The jps command lists all running Hadoop-related Java processes such as NameNode, DataNode, and ResourceManager.

```
hadoop@bmscece-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop jar Desktop/WordCount
.jar WCDriver /uzairdir/test.txt /uzairdir/out.txt
2025-04-29 15:19:10,972 INFO impl.MetricsConfig: Loaded properties from hadoop-m
etrics2.properties
2025-04-29 15:19:11,017 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot p
eriod at 10 second(s).
2025-04-29 15:19:11,017 INFO impl.MetricsSystemImpl: JobTracker metrics system s
tarted
2025-04-29 15:19:11,024 WARN impl.MetricsSystemImpl: JobTracker metrics system a
lready initialized!
2025-04-29 15:19:11,081 WARN mapreduce.JobResourceUploader: Hadoop command-line
option parsing not performed. Implement the Tool interface and execute your appl
ication with ToolRunner to remedy this.
2025-04-29 15:19:11,135 INFO mapred.FileInputFormat: Total input files to proces
s : 1
2025-04-29 15:19:11,161 INFO mapreduce.JobSubmitter: number of splits:1
2025-04-29 15:19:11,203 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_local1348329959_0001
2025-04-29 15:19:11,203 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-04-29 15:19:11,263 INFO mapreduce.Job: The url to track the job: http://loc
alhost:8080/
2025-04-29 15:19:11,264 INFO mapred.LocalJobRunner: OutputCommitter set in confi
g null
```

A MapReduce job is executed using `hadoop jar` to process `test.txt` and generate output in `out.txt`.

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -cat /uzairdir/out
.txt/part-00000
-Uzair 1
are 1
family 1
hi 1
how 3
is 2
job 1
you 1
your 2
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$
```

The output of the MapReduce job is displayed using `hadoop fs -cat`, showing the word count of the input file.

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -cat /uzairdir/tes
t.txt
hi how are you
how is your job
how is your family
-Uzair
```

The contents of the input file `test.txt` are displayed using `hadoop fs -cat`, showing a text conversation.