

# HADOOP

## 1. Starting Hadoop Cluster

```
[/]$ cd /usr/local/hadoop/sbin
[/usr/local/hadoop/sbin]$ 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 [DESKTOP-72U025C]
2020-12-08 23:10:37,345 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers
[/usr/local/hadoop/sbin]$ jps
14546 NameNode
15315 NodeManager
14948 SecondaryNameNode
15159 ResourceManager
15671 Jps
14698 DataNode
```

## 2. Creating file for counting words of

```
[~]$ mkdir hadfiles
[~]$ cd hadfiles
[~/hadfiles]$ touch wordcountfile.txt
[~/hadfiles]$ nano wordcountfile.txt
```

## 3. Moving file to Hadoop File System

```
[/usr/local/hadoop]$ hadoop fs -mkdir /programs
```

```
[/usr/local/hadoop]$ hadoop fs -copyFromLocal /home/hadoop/hadfiles/wordcountfile.txt /programs/prog1.txt
```

## 4. Running wordcount.jar, which consists of wordcount\$Map, wordcount\$Reduce, and wordcount classes

```
[~/hadfiles]$ hadoop jar /home/hadoop/hadfiles/wordcount.jar WordCount /programs/prog1.txt /programs/output/
```

## 5. Output received

```
~/hadfiles
❶ hadoop fs -ls /programs/output
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2020-12-08 23:36 /programs/output/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 69 2020-12-08 23:36 /programs/output/part-r-00000

~/hadfiles
❶ hadoop fs -cat /programs/output/part-r-00000
are 1
brother 1
family 1
hi 1
how 5
is 4
job 1
sister 1
you 1
your 4
```

## 6. Stopping Hadoop

```
~/hadfiles
❶ stop-all.sh
```