

```

package PackageDemo;
import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
import org.apache.hadoop.util.GenericOptionsParser;
public class WordCount {
public static void main(String [] args) throws Exception
{
Configuration c=new Configuration();
String[] files=new GenericOptionsParser(c,args).getRemainingArgs();
Path input=new Path(files[0]);
Path output=new Path(files[1]);
Job j=new Job(c,"wordcount");
j.setJarByClass(WordCount.class);
j.setMapperClass(MapForWordCount.class);
j.setReducerClass(ReduceForWordCount.class);
j.setOutputKeyClass(Text.class);
j.setOutputValueClass(IntWritable.class);
FileInputFormat.addInputPath(j, input);
FileOutputFormat.setOutputPath(j, output);
System.exit(j.waitForCompletion(true)?0:1);
}
public static class MapForWordCount extends Mapper<LongWritable, Text,
Text,
IntWritable>{
public void map(LongWritable key, Text value, Context con) throws
IOException, InterruptedException
{
String line = value.toString();
String[] words=line.split(" ");
for(String word: words )
{
Text outputKey = new Text(word.toUpperCase().trim());
IntWritable outputValue = new IntWritable(1);
con.write(outputKey, outputValue);
}}}
public static class ReduceForWordCount extends Reducer<Text,
IntWritable,
Text, IntWritable>{
{
public void reduce(Text word, Iterable<IntWritable> values, Context
con)
throws IOException, InterruptedException
{
int sum = 0;
for(IntWritable value : values)
{
sum += value.get();
}
con.write(word, new IntWritable(sum));
}}
}

```

**OUTPUT:**

```
pccoe@PC-07:~$ start-dfs.sh
18/08/28 17:57:21 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
Starting namenodes on [PC-07]
pccoe@pc-07's password:
PC-07: starting namenode, logging to /usr/local/hadoop/logs/hadoop-pccoe-
namenode-PC-07.out
pccoe@pc-07's password: pccoe08: ssh: Could not resolve hostname pccoe08:
Name or service not known
pc05: ssh: connect to host pc05 port 22: No route to host
PC-06: starting datanode, logging to /usr/local/hadoop/logs/hadoop-pccoe-
datanode-PC-06.out
```

```
pccoe@pc-07's password: PC-07: Permission denied, please try again.
```

```
PC-07: Permission denied, please try again.
pccoe@pc-07's password:
PC-07: starting datanode, logging to /usr/local/hadoop/logs/hadoop-pccoe-
datanode-PC-07.out
Starting secondary namenodes [0.0.0.0]
pccoe@0.0.0.0's password:
0.0.0.0: starting secondarynamenode, logging to
/usr/local/hadoop/logs/hadoop-pccoe-secondarynamenode-PC-07.out
18/08/28 17:58:21 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
```

```
pccoe@PC-07:~$ start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-pccoe-
resourcemanager-PC-07.out
pccoe08: ssh: Could not resolve hostname pccoe08: Name or service not
known
pccoe@pc-07's password: PC-06: starting nodemanager, logging to
/usr/local/hadoop/logs/yarn-pccoe-nodemanager-PC-06.out
pc05: ssh: connect to host pc05 port 22: No route to host
```

```
pccoe@pc-07's password: PC-07: Permission denied, please try again.
```

```
pccoe@pc-07's password: PC-07: Permission denied, please try again.
```

```
PC-07: Permission denied (publickey,password).
```

```
pccoe@PC-07:~$ jps
3890 Jps
3619 ResourceManager
3223 DataNode
3449 SecondaryNameNode
3034 NameNode
```

```
pccoe@PC-07:~$ hadoop fs -put wordcountFile Dest
18/08/28 18:18:18 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable

pccoe@PC-07:~$ hadoop jar MRProgramsDemo.jar PackageDemo.WordCount Dest
MRDir118/08/28 18:18:53 WARN util.NativeCodeLoader: Unable to load native-
hadoop library for your platform... using builtin-java classes where
applicable
18/08/28 18:18:54 INFO client.RMProxy: Connecting to ResourceManager at
PC-07/10.12.1.20:8032
18/08/28 18:18:55 INFO input.FileInputFormat: Total input paths to process
: 1
18/08/28 18:18:55 INFO mapreduce.JobSubmitter: number of splits:1
18/08/28 18:18:55 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1535459317953_0001
18/08/28 18:18:55 INFO impl.YarnClientImpl: Submitted application
application_1535459317953_0001
18/08/28 18:18:55 INFO mapreduce.Job: The url to track the job: http://PC-
07:8088/proxy/application_1535459317953_0001/
18/08/28 18:18:55 INFO mapreduce.Job: Running job: job_1535459317953_0001
18/08/28 18:19:01 INFO mapreduce.Job: Job job_1535459317953_0001 running
in uber mode : false
18/08/28 18:19:01 INFO mapreduce.Job: map 0% reduce 0%
18/08/28 18:19:06 INFO mapreduce.Job: map 100% reduce 0%
18/08/28 18:19:11 INFO mapreduce.Job: map 100% reduce 100%
18/08/28 18:19:13 INFO mapreduce.Job: Job job_1535459317953_0001 completed
successfully
18/08/28 18:19:13 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=138
    FILE: Number of bytes written=283721
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=176
    HDFS: Number of bytes written=32
    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
    Rack-local map tasks=1
    Total time spent by all maps in occupied slots (ms)=1699
    Total time spent by all reduces in occupied slots (ms)=1972
    Total time spent by all map tasks (ms)=1699
    Total time spent by all reduce tasks (ms)=1972
    Total vcore-milliseconds taken by all map tasks=1699
    Total vcore-milliseconds taken by all reduce tasks=1972
    Total megabyte-milliseconds taken by all map tasks=254850
    Total megabyte-milliseconds taken by all reduce tasks=295800
  Map-Reduce Framework
    Map input records=1
    Map output records=9
    Map output bytes=114
    Map output materialized bytes=138
```

```
Input split bytes=98
Combine input records=0
Combine output records=0
Reduce input groups=3
Reduce shuffle bytes=138
Reduce input records=9
Reduce output records=3
Spilled Records=18
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=71
CPU time spent (ms)=1010
Physical memory (bytes) snapshot=441659392
Virtual memory (bytes) snapshot=3845709824
Total committed heap usage (bytes)=320339968

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=78
File Output Format Counters
  Bytes Written=32
```

```
pccoe@PC-07:~$ hadoop fs -ls MRDir1
18/08/28 18:19:30 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   3 pccoe supergroup          0 2018-08-28 18:19
MRDir1/_SUCCESS
-rw-r--r--   3 pccoe supergroup        32 2018-08-28 18:19 MRDir1/part-r-
00000
```

```
pccoe@PC-07:~$ hadoop fs -cat MRDir1/part-r-00000
18/08/28 18:19:44 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
BUS    7
CAR    7
TRAIN  4
```